Author's response to reviews

Title: Indicators of sustainable capacity building for health research: analysis of four African case studies

Authors:

Imelda Bates (ibates@liv.ac.uk)
Miriam Taegtmeyer (miriamt@liverpool.ac.uk)
S Bertel Squire (sbsquire@liverpool.ac.uk)
Daniel Ansong (ansongd@yahoo.com)
Bertha Nhlema-Simwaka (bertha@africa-online.net)
Amuda Baba (amuda2b@yahoo.com)
Sally Theobald (sjt@liverpool.ac.uk)

Version: 2 Date: 19 February 2011

Author's response to reviews: see over
Responses to reviewer’s comments

Reviewer’s report
Title: Indicators of sustainable capacity building for health research: analysis of four African case studies
Version: 1 Date: 7 September 2010
Reviewer: Donald Cole

Reviewer’s report:
Below, none are major compulsory. Most are minor essential revisions with some discretionary revisions, the latter probably best decided by the editor(s).

1. Is the question posed by the authors new and well defined?
Not a question per se, but an aim [first sentence of copied section below] which is well defined

Needs more clarity on health capacity in general and health research capacity in particular. Unlike many development projects, all cases started from research studies and training in research methods was part of CSs except for CS2. The title sets up for health research. The Intro leaves it broader, but needs to be more honed, perhaps by making the link between the two but say this paper focuses on the nested health research capacity.

We have clarified the relationship between ‘health capacity’ and ‘health research capacity’ in the introduction by making the link between these two terms explicit (as the reviewer suggests) and emphasising that the focus of the paper is on health research capacity.

Similarly, capacity and capabilities seem to be used somewhat interchangeably – do you subscribe to Morgan and Baser’s distinction? If yes, can you reflect it in the article? If not, why not?

In the introduction we have indicated that ‘capacity’ and ‘capability’ are distinct (and referred to the Morgan and Baser reference) and we have now used ‘capacity’ consistently rather than ‘capability’ throughout the rest of the paper.

2. Are the methods appropriate
>> in general yes.
Selection criteria seem purposive rather than wrt set up to reduce ‘bias’. Good ones. Perhaps a supplemental file on all eleven, so know what choosing from

In order to provide enough information about each of the eleven projects for readers to understand the background, context, methods etc for each of the projects this would have to be a substantial piece of text. We are not sure that this would add value to the paper and may detract from the main message of the paper as well as substantially increasing the size of the paper.

Analysis used a prior framework – might other relevant indicators have been missed?

We have included a comment in the discussion about the potential for missing indicators by using a pre-existing framework.
and well described,
>> Yes, explicit. Good to have various authors extract, with dialogue.
and are sufficient details provided to replicate the work?
>> Hard in short space, as usual. E.g. nature of dialogue around each case, how
carried out (teleconference, in person….advantages of each?)
3. Are the data sound
>> Likely, given joint process…though independent assessment might have been
helpful to bring out other indicators e.g. currently the lead author also led the
development of the analytic framework earlier. Might some of the co-authors
brought in other elements? Or adapted the framework upon reading through the
cases.

We have included the following sentence in the discussion
‘This external and independent testing would also demonstrate whether we may have
missed any relevant indicators by using a pre-existing framework or by being closely
involved in the case studies.’

>> More detail in the narrative would help with examples, though recognize that
word count limitations may exist.

Table 1 contains details about each of the projects including the aim, capacity
building activities, context etc. We could move some of this into the narrative in the
main text if required by the editors.

and well controlled? – Not applicable

>>Should indicate that Box 1 material derived from earlier model i.e. part of
analysis approach, while Box 2 comes from analyzing cases i.e. part of results

This has been clarified in the headings for each box

4. Does the manuscript adhere to the relevant standards for reporting and data
deposition?
>> latter not relevant but former yes. Some queries, requests:
>> why call figures instead of tables?

We have changed all the figures to ‘tables’

>> Wrt presentation of indicators in Figure 2 and Box 3 (both seem to be tables
to me!), I was unclear on the need for the two tables – might they be
consolidated into one long one?

We have combined table 2 (previously figure 2) with box 3

Should there be more explicit statement of the
nature or options for indicators, particular the quantitative e.g. what would one
extract from training records and student assessments?
We have expanded some of the information in the table (now table 2) to provide more details about the nature of the indicators.

>> Could you clarify better “unique” vs “transferrable or generic” indicators.

The following statement is already in the results section – ‘Generic indicators were derived from project indicators that were used to monitor activities that were common to all case studies and relevant for sustainable capacity building. This commonality meant that they would be transferable between different projects and could be used to monitor progress towards building sustainable capacity.’

In addition we have added clarifications about unique and generic/transferable indicators where these have been mentioned elsewhere in the text.

5. Are the discussion and conclusions well balanced and adequately supported by the data?
   >> In general, yes. I struggled some between the nature of evidence presented wrt new capacity and the identification of indicators, somewhat emeshed, including with
   >> To what extent do your indicators reflect the complex capabilities (organizational as per Baser and Morgan?) cited i.e. resilience, innovation, credibility, motivation…..could you elaborate on this?

In the discussion we have elaborated on the interaction between capacity and capabilities and the fact that our indicators were focused on monitoring capacity rather than capabilities.

> Agree around resources…. Could you also extract the proportion of funds devoted to M&E, from the material of the case studies available? If not, why not, and how might this be remedied, given both your recommendation and your citing the strong role that M&E development, transferring research skills, played in some of the case studies.

We have added a statement in the results about the lack of information in the case studies about funds for monitoring and evaluation, and a statement in the discussion that emphasises the need for explicit budgeting for M&E activities.

>> Alienation of project teams wrt inflexible monitoring tools. Though I understand the problem, to base this on case studies, it would be good to draw out some of this from your case studies, or include a counter- example in which this occurred.

In the last paragraph of the discussion we have given a specific example of how this inflexibility may promote actions that are not the most appropriate for capacity development.

Conclusions
   >> Not sure your paper speaks to 1) scarcity of evidence – more like the Bates et al earlier paper or our in process systematic review.
We have removed this conclusion

>>3) some greater evidence of the changes in sophistication of indicators would be good in the results and linked to the discussion.

We have added a specific example in the results relating to stakeholders. This links to the first set of bullet points in the discussion but we have altered the wording to make this link more explicit.

>> if the evidence for the indicators is different or changes across projects, what mechanisms need to be in place to promote comparability of the evidence derived from them?

We have added a comment in the discussion about the need to regularly revise the tool and a caution about only comparing projects at similar stages of maturity

6. Do the title and abstract accurately convey what has been found?
>> note my feedback about health capacity vs health research capacity in the intro above. Same duality is in the title (health research) and the abstract (health)

Please see our response to point 1

>> factors are used in the methods section of the abstract, vs activities and evidence in the narrative (body) of the article. Using terms consistently would help.

Throughout the document we have replaced ‘factors’ with either ‘activities’ or ‘evidence’ as appropriate.

Given your assessment of the manuscript, what do you advise should be the next step?
Level of interest
- An article of outstanding merit and interest in its field (of the kind that might be found in the leading specialist journal in the field) – if their were a field of health research capacity development evaluation, this would be, given the scarcity of evidence-based work.
- An article whose findings are important to those with closely related research interests – Absolutely!
Quality of written English
7. Is the writing acceptable?
--------------------------
Please indicate the quality of language in the manuscript:
- Needs some language corrections before being published
>> A number of changes would be helpful, some words missing but could not easily indicate them, given lack of line numbers on manuscript and pdf format (or provide your reviewers with Adobe software to provide such detailed feedback!)

We have corrected typographical and grammatical errors

Declaration of competing interests
Although we have already explained in the discussion that these were projects from our own experiences, we have explicitly mentioned our involvement in these projects in the last paragraph of the introduction.

Level of interest: An article of importance in its field
Quality of written English: Needs some language corrections before being published
Statistical review: No, the manuscript does not need to be seen by a statistician.
Declaration of competing interests: I am a co-investigator with one of the authors, Imelda Bates, on a review of health research capacity development evaluation literature.
Reviewer's report

Title: Indicators of sustainable capacity building for health research: analysis of four African case studies

Version: 1 Date: 9 February 2011

Reviewer: Johan Hansen

Reviewer's report:

The paper describes the capacity building lessons to be learned from a selection of projects in Africa, fitting into an important and growing field of research. I think the paper is worth publishing as its lessons are useful to others aiming to contribute to capacity building both in Africa and other regions of the world. For one, the paper clearly shows that achieving sustainability is a long-term process that strongly depends on the ongoing commitment of those involved.

The paper contains a number of issues that need clarifying and as such can be seen as major essential revisions. By addressing these issues the paper would gain a lot in terms of providing a stronger basis for the lessons that are clearly already present in the paper itself.

For the purpose of identifying key factors to monitor capacity building, main focus of the paper is on best practices that have already achieved sustainable capacity. While these projects indeed provide valuable lessons, the design followed does not allow to determine whether the factors identified are indeed strong predictors of achieving sustainable capacity or if some of them also occur in projects that were not able to achieve sustainable capacity? It is not necessary (and probably not feasible) to alter this design, but some extra words to clarify this would contribute to the transferability of the paper’s findings. Possibly, additional references to related studies in the region could support this further.

We have added a section in the discussion to stress the importance of testing these indicators prospectively to see whether they are useful for predicting sustainability. We have not been able to find any published studies from the region that present information about indicators of sustainability for research capacity.

A related matter that would need extra words of clarification is in the selection of four out of seven case studies that best fit three different selection criteria. Of these criteria, especially the second criterion needs to be justified better, being the criterion that a case study should incorporate one of three key principles for effective capacity building (box 1). Why are there only these three principles and what is the empirical basis for this? E.g., why is it that projects need to begin with small projects designed to fill identified capacity gaps? And how does this particular principle relate to selection criterion one, namely that projects were selected that did not have capacity building as primary objective?

In addition, I would like to learn how the seven other non-selected studies fit these selection criteria: if they scored lower on criterion one and two than the four selected cases, I could understand their selection. But by also determining this selection on the second criterion it becomes difficult to determine whether the key indicators found are indeed indicators of achieving sustainability or whether they relate mostly to the criterion itself.

The three key principles for effective capacity building have been derived from the literature and the justification for these principles and their derivation have been
published in reference 6 (as cited in the text). As we have stated in the text, our case studies had to meet all of the stated criteria, not just one as understood by the reviewer. We have adjusted the wording in the methods to make this clearer.

We have expanded the information in the methods about capacity development as a secondary objective and we have added the following sentence in the discussion to explain the limitations of the study regarding the indicators of sustainability. ‘We identified indicators of sustainability retrospectively by analysing case studies which had demonstrated that they were sustainable. It will therefore be important to test whether these indicators are useful predictors of achieving sustainable capacity in a prospective long-term study…..’

When it comes to the empirical results some results are relatively clear, such as sources of project funding, while for other results it is difficult to distinguish findings from opinions. And in some cases findings appear very quantifiable, such as duration of various phases, but it is unclear based on which criteria this was determined. E.g., why is it that moment of achieving sustainable funding from the MoH is regarded as part of the expansion phase but not the consolidation phase? Would it be possible that expansion and consolidation occur simultaneously and does the framework allow for this? Some words could also be used to clarify why some studies had strongly overlapping phases, while in other cases it was possible to distinguish a clear divide between phases (see phase 3 and 4 for study 4), and again others included gaps between phases (in particular study 3): did this imply that the project stopped for some time?

We have added information in the introduction and results to explain that these phases are often overlapping and the margins of each phase are indistinct. We have also altered the titles of table 1 to indicate that these timings are approximate.

The reviewer is right in that some of the projects did stop for some time because of, for example, civil unrest in DR Congo (CS4).

Some of the unclarity is related to the manner in which the information was gathered and discussed, namely by authors themselves. Does this mean that the information was not verified by informants such as local project coordinators involved in the projects themselves? Judging from the author list, this at least does not appear to be the case in study 2 situated in Kenya.

All the authors were involved in one or more of the projects. At the time of the projects author MT was based in Kenya.

A final matter is that it is difficult to see the overlaps and differences between the large set of indicators, listed in figure 2, and a listing of examples of sources of evidence in box 3. For one, the indicators on the right hand side of figure 2 do not always appear to match the stages distinguished on the left hand side of box 3. Additionally, it is difficult to determine which of these many factors are the ones to focus on the most. Is some ranking possible, based on the wide experience of the project team, to hint which of the factors per phase are most crucial to achieve sustainability?
We have combined table 2 (previously figure 2) with box 3

**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
Reviewer's report
Title: Indicators of sustainable capacity building for health research: analysis of four African case studies
Version: 1 Date: 8 February 2011
Reviewer: Per-Olof Oستergren

Reviewer's report:
Comments to the authors:
The manuscript deals with a topic of very high current interest in its area; what predicts sustainability in programs intended to build capacity for health research in low-income (African) settings.
Since there is no given methodology in this area of research, the paper is an attempt to contribute to methodological developments as well as to provide empirical results guiding policy and implementation in the area.
The authors use three examples, all successful, for this combined exercise.
The results are very interesting, the authors arrive at suggesting both generic and specific indicators of sustainability and they present a four-stage model for the progress of successful programmes.
The paper is very convincing and could be expected to contribute to the current discussion both in the scientific community and among policy makers and implementers.

However, there are some issues which could merit some comments from the authors:
1. Since any hypothesis only could be tested if there is at least a theoretical chance that it could be refuted, why were only successful programmes selected? Could not the inclusion of programmes which failed to achieve sustainability (which should not be too hard to find…) have contributed with additional, maybe even essential information/support of the suggested theoretical framework?

We agree with the reviewer that it would be helpful to analyse the reasons why programmes are not successful in achieving sustainability. However the reasons are complex and often multi-factorial. The selection of the cases and analysis would require a different set of criteria and analytical methods from those used to select and analyse successful programmes, so we feel that this topic would be better addressed in a separate paper. We have indicated the need for such a study in the discussion.

2. What is the support of the sequence of the traits presented in the four-stage model? In this reviewer’s own (admittedly limited) experience there seem to be a possibility of an overlap of the mentioned traits. This does not necessarily invalidate the four stages, but it would be infesting if the authors commented on their view of this.

We have added information in the introduction and results to explain that these phases are often overlapping and the margins of each phase are indistinct. We have also altered the titles of table 1 to indicate that these timings are approximate.

3. Is there a possibility that the mentioned indicators could be presented in a semi-quantitative way, eg. by means of a "spider-web" type op chart, which could also give a quick visual impression of the development of the suggested indicators, and maybe present a solution of the possible problem of varying
sequential patterns of the indicators?

The indicators we have listed are only examples of the types of indicators that may be used and are therefore not an exhaustive list. We are concerned that if we present them in a diagram they may be perceived as the definitive, or perhaps the only, indicators. A spider diagram would illustrate the overlapping nature of the indicators but it would not allow us to portray the sense of increasing complexity as the projects matured.

**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.