Author's response to reviews

Title: Advancing the application of systems thinking in health: A realist evaluation of a capacity building programme for district managers

Authors:

N S Prashanth (prashanthns@iphindia.org)
Bruno Marchal (bmarchal@itg.be)
Jean Macq (jean.macq@uclouvain.be)
Narayanan Devadasan (deva@iphindia.org)
Guy Kegels (gkegels@itg.be)
Bart Criel (bcriel@itg.be)

Version: 3
Date: 25 April 2014

Author's response to reviews: see over
To,
The Editor
Health Research Policy & Systems

Sir,

Sub: Authors response to reviewers’ comments for MS: 1127865214116161 - Advancing the application of systems thinking in health: A realist evaluation of a capacity building programme for district managers

We sincerely thank the three reviewers of our manuscript for their insightful comments that have helped improve the paper. Based on their reviews, we have made changes to the manuscript. All changes have been made in track-changes mode to enable editorial review of the changes.

In this document, we have also provided a point-by-point description of how the reviewer comments have been addressed.

Please let us know in case any further clarifications are necessary.

Thanking you,
Yours truly,
N S Prashanth
Reviewer 1

Major compulsory revisions

The paper must clarify whether outcomes such as still birth rates are facility based, community based or both and the consequences of differences in the chosen sites.

We have clarified this issue in a sentence inserted on p. 12. The possible consequences of these differences for the evaluation are already discussed in the last paragraph on p. 15. A sentence has also been added to further clarify the role of these varying outcomes for our evaluation.

Make more clear that putting people up front as the most important outcome improvement factor is a choice and not a new “paradigm”.

Describe how the capacity building and training needs were defined and the people centred approach chosen. This is because of the web of causation where capacities of people is just one of the causes and may not change anything before organisational capacities are improved e.g. drugs and financing on time.

The section on p. 4 titled “Building capacity and performance” makes clear the notion of putting people at the centre of capacity-building interventions. We have added a few lines in relevant sections on p. 4 and p. 9 to further clarify the notion of people at the centre of the intervention, as well as to clarify the role of the intervention in individual and organizational capacity. The full description of the intervention itself has been presented elsewhere and references have been made to the full description in the section titled “Study setting” on p. 10.

P4 para 2 states that people and not programs change things. However, effects of programs are mentioned in the same para.

We have rewritten one of the sentences on p. 4, paragraph 2 to better reflect our position on people and programmes.

Describe more explicitly any dangers in relying mainly on two outliers. How were outlier criteria decided and does that result in a dependency bias for other criteria?

We have added a few lines under the section Limitations on p. 27-28 to clarify this.
P14 Not sure that I get the table right. The unshaded area follows a shaded one and seems to be evaluation indicators, but not all content in columns seems to correspond to overall column headings. The second shaded area is not followed by an unshaded one. Definition of rows needed.

The table has been reformatted. All shading has been removed and a new column has been added to clarify the progression from the programme input (as identified in the initial programme theory) to the context, mechanism and outcome elements identified in the refining process. A reference to the paper describing the refining of the programme theory has been added to point readers to that document.

Minor essential revisions

Training needs must surely address desired capacities of e.g. staff, but what about the real people i.e. the civil society and the population. Were their knowledge needs addressed?

The intervention that we evaluated did have a small component for training representatives from Panchayati Raj Institutions (the village-level local governments) who participate with health workers in several formal and informal platforms in the health system. We have modified a sentence under Study setting on p. 10 to reflect the facilitated discussion conducted with PRI members as a part of the intervention. In this paper, we have limited our scope to understanding the process of changes within the health services (particularly health managers). However, the possible influence of the perceived decision-spaces on organisational change has been taken into account during the refining of the programme theory in the CMO formulation (Table 1, last row).

The requirement for a defining a new theory may not be locally understood. How to go to scale?

The choice of realist evaluation (RE) as a method for this evaluation was driven by the potential for RE to be comprehensive enough to capture the various influences on capacity and performance of health workers in a health system. The resulting theory indeed has the explanatory power to explain the outcomes observed in course of the intervention. However, the ability of the theory to explain change in all district settings is certainly limited. While it may not have immediate local relevance, such insights could guide capacity-building and other human resource management initiatives at the meso level; state-level policymakers could benefit from these insights. In order to go to scale, a larger number of RE’s would be required to test and refine assertions emanating from the middle-range theory. We have alluded to this under Limitations on p. 28.

Are outcome measures attributable to the capacity building or is it really to the output and intermediate changes in management and services. Other research of more single intervention efficacy nature which show what works conditions be the main support why and output (short term and under control) is beneficial even if a direct outcome (long term and complex causation not controlled) is not possible to attribute to the program.
The evaluation seeks to establish post-hoc, the possible contribution of the intervention to the observed outcomes, using theory to guide data collection and analysis. Indeed, tracking back from the observed outcomes to intermediate workplace and individual characteristics is of proximate concern and is achieved to some degree. We are only able to propose plausible explanations on how the intervention interacted with existing conditions (either within individual participants, their workplaces or in the larger system) to produce change in some places, while not in others. Given our approach, it is not possible to make a claim that the intervention exclusively contributed to these observed outcomes. However, more iterations of refinement of observations could enable a stronger explanation for the contribution of the intervention to the observed outcomes. We have added a few sentences under the section on Limitations to explain the above.

**Discretionary revisions**

P4 para 3. individual, teams, organisational and health system. Or Individual, groups, service organisations and and the health system (CSOs and communities)?

Our conception of health system includes CSOs and communities based on the dynamic health system framework (Van Olmen et al, 2012. Analysing Health System Dynamics: A Framework). We have included CSOs and communities under the health system. By teams, we indeed refer to groups of health workers who are supposed to be working in close coordination with each other and having a shared objective (taluka health teams, district health team etc.). Given this, we wish to retain the existing dimensions for change within the health system.

Systems are complex and so are many of the figures. I takes some time to interpret them, so the main focus might be further shown in the figure or the text or some detail be merged.

We have modified figure 2 and figure 8 to improve their representation of the accompanying text.

It would be nice to have some discussion on the way to make other evaluators and actual managers use the tools at probably two levels of sophistication. Also how in future priority setting and planning may build on the same principles of factoring in adaptation to context and the main process guidance.

We have now added a new figure (figure 9) showing how the explanations (refined PTs) could be illustrated using the multipolar framework.
Reviewer 2

Major Compulory Revisions
None

Minor Essential Revisions

1. Although it is relatively clear, the paper could usefully make the distinction between realist evaluation as an evaluation tool only (assessing the how and why of an intervention related to the outcomes) versus a tool that also aims at influencing causality - ie. influencing the outcome by using the tool (see first sentence of "Conclusion" in the abstract for an example). Realist evaluation in the latter context becomes an intervention, and not merely an evaluation tool. This is sometimes a little unclear in the text, and has policy implications in that an intervention approach to the tool might have greater interest for a health manager.

2. Was the evaluation tool in any way used in an action research (or similar) approach in that health managers were given the results of the CMO analysis along the 3 year scope of the data collection period? It would be useful to clarify this in one sentence.

We have included a few lines under Discussion on p. 27 to address both these comments.

3. Following the introduction, particularly the last sentence of the first paragraph, indicating a comparison of outcomes to individual, meso, micro contextual factors, I was expecting the article to follow this outline of presenting the methodology, results and discussion. The article could benefit from a tighter structure in terms of presenting a clearer flow of thought between methods, results and discussion, either by using the stated levels of analysis mentioned in the introduction, or otherwise making it easier to follow the flow of the paper by using similar headings throughout.

We found it difficult to use the individual-meso-macro-contextual factors to structure the results as each of the talukas (cases) will then have to have separate sub-headings for these. We also found that it would make it difficult to structure our results in this format. Hence, we have used similar headings as suggested along with numbering of headings and sub-headings to enable easier navigation through the paper.

4. Page 4: Second paragraph beginning "From a realist perspective": It is mentioned that a programme is expected to work through providing "new resources" to one or more actors...... I would presume that you also mean new processes "as a new way of doing things" - not limiting change to
availability of resources only (resources being understood to be monetary, human resources, etc).
Indeed, by the introduction of new resources we anticipate the triggering of processes (a new way of doing things) that may contribute to the expected outcomes. The paragraph has been revised to reflect what we mean more accurately.

5. Page 5 first paragraph: The article states that the literature adopts research design allowing unanticipated effects etc. It would be useful with one or two examples of such designs for clarity - not only referencing them. We have clarified this by mentioning CAS thinking and theory-driven methods and their application in the conditions specified in the paragraph. Also, the next sub-section titled Realist evaluation and complexity on p. 6 further discusses the realist approach to research design in complex settings.

6. Page 8 second paragraph: Use of abbreviation PT first time without explanation.
We have included the abbreviation PT at the first instance when we have introduced the word programme theory, on p. 7.

7. Same paragraph: suggest deleting last sentence as it repeats message in sentences above.
The repetition has been removed.

8. Page 10 figure 2: Messy figure, but perhaps not easy to change as it has been described in previous paper?
We have strived to simplify the figure to the extent possible. Unfortunately, the number of actors involved and the relationships are several and hence, we were unable to simplify this further. Indeed, the figure has also been described in a previous paper and the purpose of its use here is to point the reader briefly to the structure of the intervention and referring them to the other paper for details.

9. Page 11 - Case selection, first paragraph. What are the performance indicators of the 10 talukas? Should be briefly mentioned
The section titled Case selection has been edited to reflect more accurately the process adopted. The performance indicators chosen to look at outcomes are mentioned subsequently in section 2.3 titled Data collection.

10. Page 11 - use of superscript notes: There is an extensive use of superscript and footnotes to the paper. This makes the reading of the paper complicated, particularly if important information is relegated to these notes.
11. Page 11: The particular note C should be explained in the text as it is not easy to understand the meaning of "retention of mentoring interest" as an indicator.

We have tried to provide additional information in the footnotes, especially where we felt might not be essential in the paper itself.
The retention of mentor interest is now explained in the paper itself. Footnote C has been removed.

12. Page 11 and onwards - Data Collection: The description of data collected should be considered aligned with the subscripts and footnotes to table 2. There is scope for cleaning up the description in the data collection and results section.
The alignment is now fixed and some cleaning up of the description of data collection has been done.

13. Page 12 - missing description of the "self-efficacy assessment" and "style of supervision" tool. This is also missing in table 2.
The description of the tools has been added. We have added a line indicating the publication where the full description of the tools and their application in Indian settings is discussed.

15. Page 13 - Analysis, last sentence: Incomplete use of "heuristic" term....
We have now referred to the multipolar framework and the CMO frame as heuristic tools instead of heuristic at both the above instances pointed out.

16. Page 15 - Results - second paragraph - description of figure 6: This section could possibly be moved to the discussion section, as a discussion on the variability of net change of stillbirths.
We have added a sentence at the end of this paragraph to show the relevance of these sentences in the Results section itself. Since the brief discussion on the variability of stillbirth rate has implications for the choice of cases, we have chosen to retain this (with the explanation now added) in the same section itself.

17. Page 16. Table 2: Missing superscripts 2 - 12 in the table. Consider reducing number of footnotes and including important description in the text (see note above). Avoid repetition between footnote and main text.
Since the table makes use of several terms that need expansion, the superscripts point to a legend below the table rather than footnotes.

18. Page 21: The article describes the inclusion of 2 cases (Gubbi and CN Halli). On this page there is nevertheless inclusion of data from Pavagada. This should be explained to avoid suspicion of "data mining".
The inclusion of data from Pavagada in the case summary of CN Halli is because of similarities between the two talukas in terms of their remoteness and the perceived neglect by the district expressed by the health managers. The relevant excerpts from a thematic analysis of the Pavagada case is selectively presented here, because the format does not allow a case summary of all the talukas to be presented. We have clarified this in the text added in sub-section 4.1 titled Explaining change: contribution of the intervention on p. 24.

19. Page 21: Pavagada being understaffed is repeated twice in the discussion, and could be cleaned or explained further.
The repetition has been removed.
20. Page 21 - last sentence: Justification of the term "understandably" is unclear.
This paragraph is unclear.
This paragraph has now been clarified to show how Sira taluka, while being geographically remote and socio-economically worse-off (some of the constraints that explained the frustration in the Pavagada and CN Halli) could have nonetheless shown improvement, perhaps due to the team characteristics.

21. Page 24 - first sentence: relating to the first comments, it is unclear which "institutional outcomes" this sentence is referring to.
This sentence has been modified to clarify this.

22. Page 26. It is unclear why the authors chose to use the multipolar performance assessment framework. The authors rightly mention in the Limitations section that there are a multitude of frameworks to use, and it while it is useful to choose one, the justification for the choice could usefully be included. In addition it could be useful with a brief mention of other related frameworks not used to describe impact of interventions. The introduction of the multipolar performance assessment framework could also be considered in the introduction of the article - providing a closure of analysis that would make the article easier to follow (see previous comment).
The use of the multipolar framework has now been made at the end of sub-section 1.4 titled Realist evaluation and complexity on p. 8 itself. An additional figure 9 illustrating the actual use of the framework in explaining the outcomes seen in the talukas has been added.

23. List of abbreviations - some missing? consider including PT.
We have now revised the list of abbreviations.
Reviewer 3

Major Compulsory Revisions

1. In the introduction, one Page 4, you say that “From a realist perspective, it is not programmes, but people who change things”. In fact from a Realist perspective, while agents through the exercise of agency, can change things, change occurs through the interaction of the context (the structure - to put simplistically) and the agents, and may occur, without agents exercising their agency. Frankly, in other parts of the paper this understanding is generally clear – perhaps, this statement and the parts that follow have been overlooked, and need further thinking.

Similarly, I suggest that you revisit the statement on Page 6 “In the realist view, people display behavioural tendencies innate to human beings, which manifest (or not) under specific conditions”. The phrase ‘behavioural tendencies innate to human beings’ is insufficient in capturing the Realist perspective, and is neither complete, nor accurate.

We have rewritten both the sentences to correctly reflect the realist position. We have also scanned the paper for inaccuracies in the representation of the realist position on programmes and made changes wherever, we found any errors.

2. In the methods section, page 10: Initial PT is what as you rightly point out, the explanatory pathway connecting the intervention to the expected outcomes. It is initial, and derived from the initial programme logic of the designers, literature, and actual design of the program. This is what you seek to test and refine in your evaluation. The refined PT is the product of your evaluation.

While your articulation of this idea is generally clear and consistent with how Pawson & Tilley explain it (Fig 1), the following sentences do not reflect this “Out initial PT aimed at explaining the differences in taluka outputs, accounting for differences in the individual characteristics of the health managers, institutional factors within the two taluka health services and the differing socio-political and other environmental factors. The refined programme theory of the intervention that guided the choice of data and the analysis is shown in figure 3.”

In these sentences it appears that you use the initial PT as an explanatory tool to explain the differences in ..... When in fact, you use a Realist Approach to explain, not the differences in the individual characteristics of the health managers, institutional factors, but how the differences in these characteristics/factors and their antecedents collude with the intervention to produce different effects/outputs, in similar or different contexts (Talukas).

To put simplistically, this is what a realist evaluation seeks to do. Initial PT #Realist Evaluation # Refined PT. Somewhere in the quest for articulatory sophistication, and I refer not just to the Methods section, this logic gets buried in the jargon, and the reader gets a bit lost. The ‘non realist’ reader
will benefit from a more accessible articulation. I suggest that the authors attend to this.

It is indeed the realist approach of refining the initial programme theory based on literature, programme implementation and initial data that resulted in the refined programme theory and the resulting CMO frames. We have now corrected the misrepresentation of the refining process of the PT that is expected to lead to a better understanding of the differences in outcomes, rather than the initial PT. This was an error.

We have also rewritten several portions of the paper to improve the accessibility of the methods and description of the results to the non-realist readers, especially keeping in mind the differences between the roles of the initial PT and the refined PT.

3. Section ‘Study Settings’, Page 8: I would hesitate to make such a sweeping claims that Karnataka, like rest of India, lacks a professional management cadre within the health services. Depends who you ask. The claim has a very regressive connotation; it is unnecessary. Same is the case for ‘in service training’.

We have edited the sentences to remove the regressive connotation and state only what is relevant to the study setting. The observation being made is the lack of a systematic process of in-service training or selecting managers (or people trained as managers), and rather choosing clinical specialists with several years of experience in clinical roles. We have also strengthened the observation with recent assessments and other literature that call for such a change based on positive experiences with the public health cadre in very few states in India.

4. Page 12 Data Collection, Para 3: It is not clear how all the metrics mentioned for assessment of ‘distal outputs’ of the intervention were used. Later in the Outcomes Section, it seems you used only 2 of these - ‘Utilisation Rate’ and ‘Stillbirth Rate’. This needs to be clarified.

5. It is also not clear why you choose to use Still Birth Rate as an Outcome measure when in fact you yourself rightly point out that SBR is determined by a whole range of factors. It will be good to clarify this/ reflect upon this choice better.

Although only the above two indicators are used, mention is made of the other indicators (budget utilization rate and proportion of CS performed for example in the early part of summary of cases). We have provided a justification for the use of stillbirth rate on p. 16, as well as included a paragraph on the limitations of making these purposive choice of indicators under Limitations on p. 29.

6. On Pages 17-18, you say you purposively select 2 cases to ‘illustrate how the plausible intermediate factors related to mechanisms of human, agency such as organisational commitment and self-efficacy could have contributed to the observed outcomes’. Two things here: A. This sentence, as it stands now, is not particularly meaningful to me. Please make it more accessible. B. This intent is insufficient, and neither does it capture what all you illustrate (or could illustrate) in/using these case studies. C. If however
you disagree with my reading – and indeed your intent is restricted to illustrating how only these 2 sets of mechanisms could have contributed to the observed outcomes – then I suggest you make it more explicit early on in the paper.

We have rewritten the sentences to correctly reflect our position and clarify what we mean. We also respond to the issue of other explanations and mechanisms in the section on Limitations, now modified to better reflect this limitation. Indeed, the current study studied these mechanisms based on our approach at refining the PT as well as the intent of the programme in seeking to produce change through bringing about a can-do attitude among the participants. This refining of the PT is the subject on another paper that is now referred to earlier in the sub-section 2.1 under Methods titled The realist cycle on p. 11.

7. More generally, and related to the central concept of your paper ‘capacity building’, it will be good if you unpack the concept better. As you know well, this concept has been extensively studied. For instance – if you look at how Potter and Brough (2004) look at Capacity Building, your intervention only intervenes with certain elements of the capacity pyramid. It will be good to reflect upon this situation; furthermore you might want to consider other elements as being contextual factors, more systematically and explicitly.

I think it is important to include this consideration in your analysis - irrespective of which framework you use to unpack the concept of ‘Capacity Building’.

The paper by Potter and Brough (2004) was indeed an important component in the refining of the PT, which is the subject of another paper. We have now integrated their view on capacity building as a hierarchy of needs in the sub-section 1.2 under Introduction titled Building capacity and improving performance on p. 4.

8. Similarly, at the analytical level – on Page 13 you say that you use Sicotte’s Multipolar Framework as a heuristic to explain organisational change – in your discussion section you discuss/explain your findings using this framework.

However, I would expect the Results section (pages 18 – 22 in particular) to clearly demonstrate to the reader the basis of the claims that you eventually make/points that you discuss, in the discussion section. The reader cannot clearly see the links between what you present in the Results section, the considerations in Fig 8, and the analytical claims in the discussion section. Perhaps this is clear in your mind; perhaps it is the proverbial black-box that you as the realist evaluator can see inside. What needs to happen though is that the reader should also be able to see inside this black-box too, as clearly as you as the researcher is able to do so, and arrive at the same analytical inferences. This is a common criticism that we as realist researchers must address diligently in our work, the reader shouldn’t be left to or be expected to take a leap of faith; he should on his
own also be able to see clearly inside the black-box, and clearly see the empirical basis of the realist analytical claims.
We have now moved the section introducing the multipolar framework and explained its utility in the *Introduction*. We have also revised some portions of the *Results* and *Discussion* section to address this issue.

**Minor Essential Revisions**

9. The use of the word 'insights'. Insight is a non-countable 'noun'.
This has been corrected in all the instances where we used *insights*.

10. Page 11: “In a second step, cases are selected purposively”. Please ensure that you consistently use the same tense throughout the paper.
We have reviewed the usage of different tenses in the manuscript and revised incorrect instances.

11. Page 17: the phrase ‘negligent change’ is wrong, perhaps you meant ‘negligible change’.
This has been corrected.

**Discretionary Revisions**

12. It will be good if the Figures are made snappier. Some consistency in style and form should help.