

## PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (<http://bmjopen.bmj.com/site/about/resources/checklist.pdf>) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

### ARTICLE DETAILS

<b>TITLE (PROVISIONAL)</b>	Improving the performance of community health workers in humanitarian emergencies: A realist evaluation protocol for the PIECES programme.
<b>AUTHORS</b>	Gilmore, Brynne; Adams, Ben; Bartoloni, Alex; Alhaydar, Bana; McAuliffe, Eilish; Raven, Joanna; Taegtmeier, Miriam; Vallières, Frédérique

### VERSION 1 - REVIEW

<b>REVIEWER</b>	Bruno Marchal Department of Public Health, Institute of Tropical Medicine, Antwerp
<b>REVIEW RETURNED</b>	04-Apr-2016

<b>GENERAL COMMENTS</b>	<p>Peer review Improving the performance of community health workers in humanitarian emergencies: A realist evaluation protocol for the IMPaCT programme. bmjopen-2016-011753</p> <p>General comments This paper is well written and merits publication if the authors address the comments below, which mainly deal with the realist methodology.</p> <p>Major changes</p> <p>P10, line 22-23: "This study employs a mixed-methods realist evaluation using multiple case studies within the IMC CHW programme." To avoid confusion, I would suggest to delete 'mixed-methods'. The main methodology adopted by this study is RE. It is the data collection methods that will be used - combining quantitative and qualitative data collection methods – that make this a study using mixed methods, which is in line with realist principles. However, it is not just a mixed methods study, which arguably are often a-theoretical and not theory-driven.</p> <p>P10, line 29-30 : "As humanitarian emergency contexts most likely have unique factors that differentially contribute to CHW performance and motivation..." Realists would hold that context matters, but not necessarily that contexts have unique features. Rather, there are patterns (demi-regularities) to be found, not only described by in CMO configurations but also in terms of typologies of contexts...</p>
-------------------------	---

	<p>P11, line 12-13: “Realist evaluations have a similar methodological process to other more traditional research approaches, with their iterative nature allowing for findings to be further tested and continually refined.” Here, further explanation of the similar methodological processes, the traditional research approaches and the iterative nature needs to be given.</p> <p>P11, line 21-28: “Mixed-methods were selected in order to (i) measure the various latent constructs described in the IPT, (ii) ascertain the association between these latent constructs, and (iii) understand how these variables influence one another to impact on CHW performance”</p> <p>This needs further explanation: which methods are going to be used for which objective?</p> <p>P11, line : “Case studies are purposefully selected to fit with the initial programme theory with intra-programme (same programme run across different groups) case studies being ideal to refine and develop further programme theories and increase the transferability of the findings” This needs some revision. First, not case studies are selected, but ‘sites’ where the ‘case’ can be studied (see Yin R: Case study research. Design and methods, Third edn. London: Sage Publications; 2003). Second, in RE, the case selection is purposive, not in order to have cases that fit with the initial PT, but rather to ensure that the cases provide sufficient opportunities to ‘test’ (parts of the) initial PT. Third, what is called ‘intra-programme case studies’ that focus on the ‘same programme rune across different groups’ is not necessarily the best option: same programme-different outcomes, or same programme-different context can be as interesting.</p> <p>Table 1 - Literature and Documentation It would be good to insert the references to the original sources of the theories mentioned in this table.</p> <p>P12, line 37-38 It would be good to mention here that you will adopt for this particular study the multiple case study design (see for instance Yin 2003).</p> <p>P13, line 5-6: “... validated scales measuring the variables ...” Realists aim at explaining outcomes. In that sense, it would be good to explain how the proposed measurements will help determining which CMOs hold to explain the observed outcomes. In other words, it is about measuring the indicator(s) for each mechanism and assessing the associations between intervention, ‘variables’ and observed outcomes. It would be more clear and in line with RE analysis if the term ‘variable’ is replaced by ‘indicator for mechanism’ or a similar expression.</p> <p>P15, line 43: “Stakeholders (actors) inherently have mechanisms and these, combined with the right contextual factors, produce the generative mechanisms related to outcomes [58].” This is quite condensed an explanation of the CMO approach. Please unpack and clarify what you mean by this.</p> <p>Minor changes</p>
--	--

	<p>P10, line 3: "... CHW performance, will further elucidate the specific needs of CHWs in specific contexts ...": Please clarify which needs these are and in how far these are related to management or other interventions.</p> <p>P10, line 38: van Belle: correct to Van Belle</p> <p>P10, line 40-41: "Are realist evaluation sees programmes as theories incarnate the IPT describes how..." Correct: As realist</p> <p>Figure 2:</p> <ul style="list-style-type: none"> <li>• 'ti' was replaced by '=' and needs correction... and some columns are obscured.</li> <li>• The meaning of the dotted arrows (from right to left) on top between intervention, mechanism and outcome would indicate what exactly?</li> </ul> <p>P15, line 54-55: "Once the initial analysis in each context is complete"</p> <p>In each context, or of each case? Each case may present you with differences in context (over time)...</p> <p>P15, line 56: "The finalised refined programme theory (the mid-range theory)..."</p> <p>I would argue that even if PT and MRT are used interchangeably in the current RE literature, the refined programme theory and MRT do not operate at the same level/ MRTs are middle range theories as you indicated (Merton's definition) and are more abstract than programme theories. I would say that you ill refine the initial PT through each case study, and that the comparison of the refined PTs may help you formulate an MRT that is of the middle range...</p> <p>General remarks</p> <p>P14 Analysis</p> <p>I would just like to stress that the realist analysis is all about the (intervention)-(actor)-context-mechanism-outcome configuration and that the findings of the multiple mediation conditional process analysis indeed need to be framed within the results of additional data analysis to go attain a configurational approach to analysis. Realist analysis is case-oriented, not variable-oriented analysis</p>
--	---

<b>REVIEWER</b>	Kevin Pottie University of Ottawa, Canada
<b>REVIEW RETURNED</b>	19-Apr-2016

<b>GENERAL COMMENTS</b>	<p>Thank you for the opportunity to review this very interesting and ambitious protocol. It takes on an important and under researched are, CHW in humanitarian emergencies. Specifically it focused on Level III Emergencies. The study is focused on understanding what enhances the motivation and performance of community health workers.</p> <p>Mixed methods realist evaluation across 3 countries, Turkey, Iraq, Lebanon. Based on interviews, life histories, and critical incident narratives, and variable modelling of key constructs to explain how contextual factors trigger mechanisms for outcomes.</p> <p>Comments Concerns and Suggestions:</p>
-------------------------	--

	<p>The review plans for considerable field data collection in 3 different humanitarian settings, although the team may be well equipped to take on this field work, feasibility may be an issue. There may be a need for realist review flexibility for this protocol.</p> <p>I wondered why in their methods they would not have used some of the existing literature and then supplement this with the field data? It appears they will use it in terms of theory, constructs and modelling, but this is not really clear how this is happening. More clarity would be good.</p> <p>They do a very good job documenting relevant CHW literature and using it as background. But they could provide a bit more details on why this literature would be different from the actual humanitarian settings. This could provide a more focused foundation for the review and it could increase feasibility. This could help mitigate challenges when taking on challenges of humanitarian settings.</p> <p>If they used previous publications the format of the protocol could be more aligned to the RAMSES reporting guideline for realist reviews. Authors should explain why they chose not to review to go in this direction. Within the realist world, there has been some debate as to how a realist review protocol can best be constructed, given the high degree of flexibility that is needed. More mention of how this review will address challenges and need for flexibility would be most helpful. I am happy to see this protocol emerge and this is the direction, protocols, I would like to see thing go.</p> <p><b>Impression-</b> good protocol, important subject, but some revisions would be needed- more details on how pas resources can be used to guide data collection and analysis, discuss why they did not use a RAMSES methods approach and better report on feasibility and flexibility plans.</p>
--	--

### VERSION 1 – AUTHOR RESPONSE

In response to Bruno Marchal’s comments we have made the following amendments:

P10, line 22-23: To avoid confusion, I would suggest deleting ‘mixed-methods’.

Revisions Made:We have removed mixed-methods and changed this to, “This study employs a realist evaluation using multiple cases across purposely-selected humanitarian emergency contexts where IMC is currently implementing CHW programmes”

P10, line 29-30: Realists would hold that context matters, but not necessarily that contexts have unique features. Rather, there are patterns to be found, not only described by CMO configurations but also in terms of typologies of contexts.

Revisions Made:We have changed this sentence to “The complexity and variability of CHW programmes across settings lends itself particularly well to realist studies, and realist methods have been recently recommended for the study and understanding of CHW motivation and performance. Realist methods are particularly applicable to humanitarian emergencies, where CHW motivation and performance are likely to show a different pattern”.

P11, line 12-13: A further explanation of the similar methodological processes, the traditional research approaches and the iterative nature needs to be given.

Revisions Made:This has been elaborated upon as follows: “As stated by Pawson and Tilley, realist

evaluation research follows the 'traditional research cycle' of hypothesis (theory) generating and subsequent testing through data collection and analysis<sup>53</sup>. Realist evaluations encourage iterative data collection and/or the use of concurrent case studies to provide findings for increased programme specification.

P11, line 21-28: "Mixed-methods were selected in order to (i) measure the various latent constructs described in the IPT, (ii) ascertain the association between these latent constructs, and (iii) understand how these variables influence one another to impact on CHW performance". This needs further explanation: which methods are going to be used for which objective?

Revisions Made: Further clarification is provided to outline which objectives are addressed by which methods. This section now reads:

In the case of this particular study, a combination of quantitative and qualitative methods was selected. Specifically, quantitative methods will be used to (i) measure the various latent constructs described in the IPT using confirmatory factor analysis, (ii) ascertain the association between these latent constructs using multiple mediation analysis. Qualitatively, life histories and critical incident reviews with CHWs will be used to understand how these variables influence one another to impact on CHW performance.

P11, line : "Case studies are purposefully selected to fit with the initial programme theory with intra-programme (same programme run across different groups) case studies being ideal to refine and develop further programme theories and increase the transferability of the findings" This needs some revision. First, not case studies are selected, but 'sites' where the 'case' can be studied (see Yin R: Case study research. Design and methods, Third edn. London: Sage Publications; 2003). Second, in RE, the case selection is purposive, not in order to have cases that fit with the initial PT, but rather to ensure that the cases provide sufficient opportunities to 'test' (parts of the) initial PT. Third, what is called 'intra-programme case studies' that focus on the 'same programme run across different groups' is not necessarily the best option: same programme-different outcomes, or same programme-different context can be as interesting.

Revisions Made: This section has been rewritten and now reads: Research sites were purposefully selected to ensure that they provide sufficient opportunities to test parts of the IPT. Intra-programme studies (e.g. the same programme implemented across different groups) or same programme being run in different contexts are useful to refine and develop further programme theories and increase the transferability of the findings [53].

#### Table 1 - Literature and Documentation

It would be good to insert the references to the original sources of the theories mentioned in this table.

Revisions Made: We have added references where the table first mentions key constructs and theories (i.e. self-determination theory, burnout, organizational commitment, performance, etc.)

P12, line 37-38 It would be good to mention here that you will adopt for this particular study the multiple case study design (see for instance Yin 2003).

Revisions Made: This has been added and the sentence now reads: For this particular study, we will adopt a multiple case study design <sup>55</sup>, whereby the IPT will be further refined within each case study (i.e. Iraq, Lebanon, and Turkey) to develop a more contextually relevant programme theories of how to enhance CHW performance in humanitarian settings.

P13, line 5-6: "... validated scales measuring the variables ..."

Realists aim at explaining outcomes. In that sense, it would be good to explain how the proposed measurements will help determining which CMOs hold to explain the observed outcomes. In other words, it is about measuring the indicator(s) for each mechanism and assessing the associations between intervention, 'variables' and observed outcomes. It would be more clear and in line with RE analysis if the term 'variable' is replaced by 'indicator for mechanism' or a similar expression.

Revisions Made: The term 'variables' has been replaced with 'indicators for mechanism'. This now reads: "First, and aligned to the International Test Commission principles<sup>56</sup>, existing, and where possible, validated scales measuring the indicators for the mechanisms outlined in the IPT will be translated and back-translated. These indicators for mechanism will then be measured at three different time points throughout the proposed 24-month study: baseline (month 3), midterm (month 12) and endline (month 21)"

P15, line 43: "Stakeholders (actors) inherently have mechanisms and these, combined with the right contextual factors, produce the generative mechanisms related to outcomes [58]."

This is quite condensed an explanation of the CMO approach. Please unpack and clarify what you mean by this.

Revisions Made: We have expanded and now state: "Using realist evaluation generative succession logic, underlying mechanisms (M) associated with the programme are triggered, or not, by or under particular contextual factors (C) or programme conditions. This interaction between generative Ms and Cs produce outcome patterns (O) of the programme.

P10, line 3: "... CHW performance, will further elucidate the specific needs of CHWs in specific contexts ...": Please clarify which needs these are and in how far these are related to management or other interventions.

Revisions Made: We have elaborated that these are personal and/or professional needs that the CHW may require to be fulfilled for improved performance, arising factors from our IPT, and highlighted that these learnings may be applied to similar intervention programmes.

P10, line 38: van Belle: correct to Van Belle Revisions Made: Corrected

P10, line 40-41: "Are realist evaluation sees programmes as theories incarnate the IPT describes how..." Correct: As realist Revisions Made: Corrected

Figure 2:

- 'ti' was replaced by '≡' and needs correction... and some columns are obscured.
- The meaning of the dotted arrows (from right to left) on top between intervention, mechanism and outcome would indicate what exactly? Revisions Made: Figure 2 has been corrected

The following caption has been added to Figure 2:

Figure 2- PIECES' Initial Programme Theory. The bidirectional errors depict that change (or producing an outcome) does not necessarily happen in a unilateral direction. The Context, Mechanisms and Outcome all influence one another.

P15, line 54-55: "Once the initial analysis in each context is complete"

In each context, or of each case? Each case may present you with differences in context (over time)...

Revisions Made:Corrected. Now reads: Once the initial analysis of each case is complete, an intra-programme case study comparison will occur.

P15, line 56: "The finalised refined programme theory (the mid-range theory)..."

I would argue that even if PT and MRT are used interchangeably in the current RE literature, the refined programme theory and MRT do not operate at the same level/ MRTs are middle range theories as you indicated (Merton's definition) and are more abstract than programme theories. I would say that you ill refine the initial PT through each case study, and that the comparison of the refined PTs may help you formulate an MRT that is of the middle range...

Revisions Made:Corrected. Now reads: The IPT will subsequently be refined through each case study, and the comparison of the refined programme theory may help formulate a programme theory that is of middle range.

• We have also corrected this throughout the remainder of the document, making clear that we are working to refine programme theories not MRT (which as stated earlier may in turn be used to help formulate a middle-range theory).

In response to Kevin Pottie's comments we have made the following amendments:

#### K. Pottie's Comments

The review plans for considerable field data collection in 3 different humanitarian settings, although the team may be well equipped to take on this field work, feasibility may be an issue. There may be a need for realist review flexibility for this protocol.

Revisions Made: Added on pg. 17: "Additionally, working across three complex humanitarian settings may present unforeseen challenges and requires a level of flexibility and/or adaptability of the protocol during the research process. Any alterations to the schedule will be discussed with the research consortium and efforts will be made to maintain the integrity of the overall protocol, with rigour and realist evaluation standards kept as a priority. If the level of flexibility required to continue the research deviates too far from the protocol, the consortium will evaluate the ethical and methodological consequences before agreeing on a best way forward".

I wondered why in their methods they would not have used some of the existing literature and then supplement this with the field data? It appears they will use it in terms of theory, constructs and modelling, but this is not really clear how this is happening. More clarity would be good.

Revisions Made: We appreciate the need for clarity on this point. We believe you are referring to how we developed the initial theory (which was a combination of literature and field data through the FGDs and IDIs). We did use the literature to guide the field data collection tools, with an analysis of the literature and the data informing the initial programme theory (which then informed the overall study protocol). We have added more clarity to this point on pg. 11: "Important to note is that document collection and analysis was done prior to stakeholder input, with context-mechanism-outcome configurations (CMOCs) being developed and further refined through the FGDs and SSIs. The interview guides were developed based on literature/document findings. This process was done to provide contextual information for further refinement and elicitation of the initial programme theory."

They do a very good job documenting relevant CHW literature and using it as background. But they could provide a bit more details on why this literature would be different from the actual humanitarian settings. This could provide a more focused foundation for the review and it could increase feasibility. This could help mitigate challenges when taking on challenges of humanitarian settings.

Revisions Made: On page 5, we offer the following reasons for why the CHW performance literature differs in Level III emergencies. Namely, "Health services in humanitarian emergencies are frequently non-existent or under pressure because of on-going violence and conflict 35 yet the needs for health care are increased. The impact of humanitarian emergencies on a population's health is severe and exacerbated by increases in food insecurity, population displacement, crowding and poor access to water and sanitation, lack of resistance to infection, the physical and psychological effects of weapons and exposure to violence, and the collapse of basic health care services 36. The impact of humanitarian emergencies on health workers and service provision is also extensive and includes the destruction of health facilities, infrastructure, frequent and prolonged shortages in drugs and equipment, loss of qualified health staff, and restricted access to healthcare 37."

If they used previous publications the format of the protocol could be more aligned to the RAMSES reporting guideline for realist reviews. Authors should explain why they chose not to review to go in this direction.

Revisions Made: We chose to format our protocol publication loosely on other realist evaluation

protocols that have been published in BMJ Open, as well as following the format of the Realist Evaluation cycle to provide an overall guide and logic-flow. Though the RAMESES reporting guidelines for realist reviews/syntheses were consulted for aspects of the paper, there are currently no formalised guidelines for reporting realist evaluations and/or their protocols.

Within the realist world, there has been some debate as to how a realist review protocol can best be constructed, given the high degree of flexibility that is needed. More mention of how this review will address challenges and need for flexibility would be most helpful.

Revisions Made: Thank you for this point. Please see our response from your first comment and additions on pg. 17, which we believe covers this point as well.

#### VERSION 2 – REVIEW

<b>REVIEWER</b>	Bruno Marchal Institute of Tropical Medicine, Antwerp
<b>REVIEW RETURNED</b>	18-May-2016

<b>GENERAL COMMENTS</b>	In this revision, the authors addressed my concerns satisfactorily and thus I recommend publication.
-------------------------	--