Reviewer's report

Title: What are the barriers and facilitators for third sector organisations (non-profits) to evaluate their services? A systematic review

Version: 0 Date: 04 Dec 2017

Reviewer: Rob Macmillan

Reviewer's report:

Thank you for submitting the paper. I am pleased to have been invited to review it, and have found it interesting and enjoyable to read. Please accept my apologies for the delay in providing a report.

Overall assessment:

This is a very good paper. It is mostly well-written, straightforward and clearly presented and structured. I learnt a lot through reviewing it, and, if it is published, the authors will have done a very useful job of assessing literature in this otherwise fragmented area. It provides a helpful and up to date ground-clearing service, and, sets the scene well for further research. Subject to the comments below, I believe it is publishable, but have four main suggestions designed to improve the paper - though I should stress: I do not think this requires a major re-write. In addition, I have made a more substantial number of (usually very) minor suggestions as a result of a close read of the paper, which are listed in bullet point form at the end of this review.

The four main things I would suggest the authors attend to in a revised submission are:

1. Note and utilise the distinction between 'capacity' and 'capability'

The phrase 'lack of expertise and internal capacity' (p.13, lines 327-329) involves, strictly speaking, two separate concepts. The former is about 'capability' (a reference to skill or expertise - 'know-how', or how well something can be done) whereas the latter is 'capacity' (a volume conception of how much or how little something can be done). 'Lack of financial resources' is a capacity issue rather than a capability issue (as is 'lack of time' - p.14, line 341). At the risk of self-referencing, the distinction between them is discussed in Macmillan, R. and Ellis-Paine, A., with Kara, H., Dayson, C., Sanderson, E. and Wells, P. (2014) Building Capabilities in the Voluntary Sector: What the evidence tells us (TSRC Research Report 125, Birmingham, Third Sector Research Centre)(see, e.g. pp.6-7), informed by the Big Lottery Fund's thinking in this area (see: https://www.biglotteryfund.org.uk/research/making-the-most-of-funding/building-capacity). The distinction is well represented in the two main thematic groupings subsequently used in the paper, i.e. 'Factors related to lack of resources' refers to lack of capacity, whilst
'Factors related to technical capacity and evaluation skills' refers to 'capability.' The suggestion is to make use of this distinction in discussing the analysis and results of the review.

2. Provide more information on the assessment of study quality

It is not clear on what basis studies were assessed as being of 'high', 'medium' or 'low' quality (p.10, line 244). The checklists in Appx A and B are helpful, but it would be helpful to say a little more about this in the text (see also p.13, lines 309-316, where reference is made to 'achieving a score' - what does this mean?).

3. Modify the arguments supporting the promotion of EBP in the discussion and concluding sections

The rationale for promoting EBP seems a little overplayed and, I thought, stretched beyond the results of the review. I felt that the final paragraph on p.18 (and through to p.19) involves something of an unwarranted translation: a reported problem with indicators ('Challenges in identifying accepted outcome and impact indicators') is turned into a need for impact evaluation frameworks and guidelines. This is not, it seems to me, to be what is reported as a barrier, nor indeed as a facilitator. Likewise, do we know that '…consensus procedures from EBP remain largely unused by the third sector…' (p.19, line 485-6)? This reads like an over-claim, and is either in need of a supportive reference or argument, or, I would suggest, should be toned down. Finally, there is a need for care to avoid slippage in the reference to '….may reinforce poor evaluation practice among TSOs' (p.19, line 472). The review does not provide evidence of poor evaluation practice, only of reported and interpreted barriers and facilitators.

4. Necessary and sufficient conditions

The paper concludes with the argument that '…it is central not to focus on individual barriers (such as financial resources), but to reach consensus, or at least have a procedure for consensus, to determine what evaluation criteria different types of TSOs require…' (p.22, lines 544-555). The aim of the argument here seems to be to draw attention away from financial barriers and towards evaluation criteria. This does not seem to be a justifiable step when expressed in these terms. For example, financial barriers seemed actually to win out in the 'vote count' against 'identifying accepted outcome and impact indicators' (14 against 13). Rather than overlook financial barriers, as the present wording of the text suggests (e.g. 'it is central not to focus on…'), my suggestion would be to distinguish between 'necessary' and 'sufficient' conditions for improved evaluation practice in the third sector. My reading of the systematic review is that overcoming financial and similar 'capacity' barriers would be necessary, but not sufficient, to improve practice (given the existence/prevalence of other factors), but, equally, securing a consensus on evaluation criteria (and other 'capability' issues) is also necessary, but not sufficient
(given the significance of financial and other capacity constraints). Hence the review overall might conclude that 'capacity' and 'capability' (in this case in evaluation practice) are separate but very closely related.

Minor notes, queries, typos, etc.

- Title: do you really 'experience' a facilitator? Can you think of another word, e.g. 'encounter'? or just 'What are the barriers and facilitators for good evaluation practice in third sector organisations? A systematic review'

-p.2, lines 31-32 (abstract): is it true to say that 'many TSOs fail to evaluate their activities following the principles of evidence based practice?' how do we know? The review is of reported barriers and facilitators, not of actual evaluation practice.

-p.4, lines 82-83: the cited definition of EBP would seem to be irrelevant to many TSOs, given the reference to 'clinical expertise' and 'patients'. I'd suggest using a broader or more generic definition.

-p.4, line 89: '…community-based organisations…' (plural)

-p.4, lines 92-95: this definition of TSOs was generated by the Johns Hopkins Comparative Nonprofit Sector project, so perhaps the reference should be to the original. In a UK setting, the reference is Kendall, J. and Knapp, M. (1996) The Voluntary Sector in the UK (Manchester, Manchester University Press)

-p.4, lines 101-102: I don't think it is true that 'most' TSOs in the UK focus on delivering 'social services' Does the Almanac data really show this?

-p.5, line 103: '…annual spending of £45.5bn…' should be 'annual income (spending was £43.3bn)

-p.5, lines 104-106: 'Of these budgets…' is confusing, as it points attention to the two figures in the preceding sentence. But the government's £15.3bn funding to the sector is (a) greater than the two figures in the preceding sentence, and thus (b) should relate to the overall sector income figure, of £45.5bn
- p.5, line 112: after '…their sample…' suggest adding in what the sample figure was, i.e. 8% of what?

- p.5, line 112: a reference is given here to 'Breckell et al 2010', but this seems out of line with the referencing approach taken in the rest of the paper

- p.6, lines 140-141: '…This scoping review, which considered neither study quality nor….'

- p.6, lines 145-149: I'd suggest rephrasing the passage here, as currently it sounds like it was remiss of Hardwick et al not to have considered evaluation processes, rather than simply that this was not their chosen focus.

- p.8: were there any time restrictions used in the search strategy? (i.e. 'material published from 2000 onwards')

- p.9, line 218: delete 'is'

- p.9, line 219-220: the statement '…studies were either qualitative, mixed methods or cross-sectional…' does not seem to make sense, as the three categories here are not mutually exclusive. Quantitative research can be either longitudinal or cross-sectional, but so can qualitative research. The paper seems to imply that 'cross-sectional' studies are automatically quantitative (see also p.10, line 235). Maybe it would be better to say 'qualitative, survey-based or mixed methods…'

- p.10, line 234: It isn't clear why the variability of numbers of TSOs in the studies is a reason for undertaking a quality appraisal. I'm not sure a reason is needed - assessing quality is an important step in its own right.

- p.10, line 253: five studies were excluded because they were '…about community participatory research processes…' It isn't clear what this means and why they were excluded - an additional explanation is needed.

- p.11, line 265: full text needed for 'LMIC' before acronym

- p.11, line 270: Suggest sub-title as 'Samples of TSOs in included studies'
-p.11, line 275: '….a 1000 or more TSOs…' rather than '…more than 1000 TSOs…'

-p.12, line 282: the bias towards 'larger organisations' needs an explanation or definition, since 'large' and 'small' in the context of TSOs are used vaguely and inconsistently - are the income bands used in the subsequent sentence what you mean here?

-p.12, line 283: add 'annual income' before '….below £100,000'

-p.12, lines 284, 285: '£1million' not '1£million'

-p.12, line 291: 'in' rather than 'with' social and human services?

-p.13, line 314-315: suggest adding 'surveys' before '…of a cross-sectional nature…' and then remove 'survey' from '…validate the survey findings…'

-p.13, line 329: 'appropriate evaluation goals (15/24)' should be '…(16/24)…' if table 2 is correct.

-p.14, line 335: is table 1 really needed, given that it is fully described in the preceding text? I'd suggest one or the other is needed but not both.

-p.14, line 340: suggest removing 'survey' before 'respondents', as the findings in some of the studies do not come from surveys

-p.14, lines 343-344: 'supports' and 'requires', since 'management' is singular

-p.15, line 372: '…were those…' rather than '…was factors…'

-p.15, lines 380-382: I found this sentence to be very unclear: can it be reworded?

-p.16, line 384: suggest '…only two studies, both qualitative…'

-p.16, line 391: because it isn't immediately obvious what a 'sensitivity analysis' entails, I'd suggest adding 'by' between 'conducted' and 'including'
-p.17, line 415: suggest 'for example' rather than 'e.g.' here (and in subsequent places)

-p.17, line 416: '…appropriate support…' rather than '…the appropriate support…'

-p.17, line 418: I'm not sure what a 'commissioning contract' is. I'd suggest a more general reference - 'grants or contracts'

-p.17, line 422: 'procedures'

-p.17, line 430: suggest adding 'and facilitators' after 'external barriers', since the preceding paragraph is about both barriers and facilitators

-p.17, line 432: move 'both' to before 'barriers'

-p.18, line 456: suggest 'The most cited group of factors…' rather than 'One of the most cited themes of factors…'

-p.18: is it worth a line or two exploring why 'lack of resources' features as a barrier, but 'more resources' doesn't feature so much as a facilitator?

-p.18, line 459: it would be important to recognise that 'lack of resources' was cited by just as many times as lack of evaluation skills.

-p.19, line 466: 'social impact bonds' are not evaluation tools or frameworks; they are financial mechanisms. Also it should be 'Social Return on Investment' not '…of investment'

-p.19, line 469: typo: '…no any…'

-p.20, line 509: suggest using 'fields' or 'types of organisation' rather than 'sectors' here, since 'sector' is also used typically as a label for the aggregate of voluntary, community and non-profit organisations

-p.21, lines 522-523: '…most third sector spending originates from Western TSOs…' Is this true? A reference is needed to support this claim.
-p.21, line 533: You might also add Jenny Harlock’s TSRC Working Paper from June 2014 (No.123) to reference 87, as it reports commissioners’ views on evaluation requirements

-p.22, line 552: '….are put at risk…. ' (not 'is')

-p.22, line 559: suggest '…to assess with confidence…' rather than 'certainty', which seems like an impossibly high threshold.

-pp.24-32 References: some of the references are not fully cited. For example, TSRC working papers have details missing: e.g. Reference No.4 (page 24) would be 'Harlock, J. (2013) 'Impact measurement practice in the UK third sector: a review of emerging evidence' TSRC Working Paper 106, July 2013 (Birmingham, University of Birmingham), followed by a direct weblink?; References 28, 32 and 33 seem to be of the same report?; References 1 and 75 are to the same item.

Level of interest
Please indicate how interesting you found the manuscript:

An article whose findings are important to those with closely related research interests

Quality of written English
Please indicate the quality of language in the manuscript:

Needs some language corrections before being published
**Declaration of competing interests**

Please complete a declaration of competing interests, considering the following questions:

1. Have you in the past five years received reimbursements, fees, funding, or salary from an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

2. Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

3. Do you hold or are you currently applying for any patents relating to the content of the manuscript?

4. Have you received reimbursements, fees, funding, or salary from an organization that holds or has applied for patents relating to the content of the manuscript?

5. Do you have any other financial competing interests?

6. Do you have any non-financial competing interests in relation to this paper?

If you can answer no to all of the above, write 'I declare that I have no competing interests' below. If your reply is yes to any, please give details below.

For a short time one of the authors (PM) and I were employed in the same school at Birmingham - after he joined from Oxford and before I moved to Sheffield Hallam. However, we were never introduced, and I have completed the review whilst at my new institution. There are no competing interests

I agree to the open peer review policy of the journal. I understand that my name will be included on my report to the authors and, if the manuscript is accepted for publication, my named report including any attachments I upload will be posted on the website along with the authors' responses. I agree for my report to be made available under an Open Access Creative Commons CC-BY license (http://creativecommons.org/licenses/by/4.0/). I understand that any comments which I do not wish to be included in my named report can be included as confidential comments to the editors, which will not be published.

I agree to the open peer review policy of the journal