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

TITLE (PROVISIONAL)	A qualitative study of self-evaluation of Junior Doctor performance: Is perceived 'safeness' a more useful metric than confidence and competence?
AUTHORS	Roland, Damian; Matheson, David; Coats, Timothy; Martin, Graham

VERSION 1 - REVIEW

REVIEWER	Nai Ming Lai School of Medicine, Taylor's University, Malaysia.
REVIEW RETURNED	09-Jun-2015

GENERAL COMMENTS	Thank you for the opportunity to share my view on this interesting manuscript. In this study, the authors focused on the confusion that exists between the definitions of confidence versus competence, and gathered a group of junior doctors in a structured and in-depth discussion to try to make sense of these two terms using the management of febrile children as the context. They identified the concept of "safeness" as the key concept that is common in all discussion relating to confidence or competence, and they suggested that such term as "safeness" be used in preference to the terms of confidence and competence in any evaluation of clinical ability. The authors have done well in demonstrating the inadequacy of the terms "confidence" and "competence" when used to illustrate self-perception of ability, and they went to great lengths to decipher the messages generated following intensive discussions among the participants. However, I feel that there are three issues that have limited the utility of the study. First, the authors gave the impression that the terms "confidence" and "competence" were both descriptions of self-perceived ability, and the existing confusion between the two terms was resolvable by separating the two terms into descriptions on different aspects of self-perceived ability. One important aspect that they might have overlooked in their attempt to make sense of these terms was the assessor. Confidence is often cited in the context of self-perception while competence is often used when assessment is performed by another person. In this regard, the methodology of the current study was limited by the absence of views from other stakeholders in the determination of a doctor's ability, most notably the patients. The second issue concerns the ultimate validity of the terms. A major difficulty in defining a term that denotes clinical ability lies in the fact that any such term is only a surrogate measure of the actual patient outcome. Often, in the absence of such outcome data as the reference standards in the same study, it is difficult to validate each measure. This study had the same predicament, as we were unclear

	whether “safeness”, as quoted by the authors, were a better concept to denote clinical ability compared to confidence and competence, as there was no objective validation to this term in the form of direct evaluation of the degree of “safeness” against patient outcomes, for example, the rate of mortality or missed diagnosis. Additionally, this study was conducted using a specific context in managing febrile children. As implied by the authors in their discussion, this placed a limitation on the generalisability of the findings. I would imagine that “safeness” as a term might not be adequate in defining the ability of a clinician in situations like patient explanation or counselling and effective organisation of team work from acute resuscitation to multidisciplinary care of chronic problems. I feel that it is important that the authors clearly highlight these issues as limitations, and confine the concept of “safeness” derived from the study to educating young medical trainees in helping them better understand how they see their own clinical abilities through reflection, rather than introducing the term as an actual measure of ability as an alternative to confidence or competence. Following are some suggestions pertaining to specific parts of the paper: Introduction, paragraph 1 “An accurate and valid measure of confidence and competence is important because the least competent professionals are also the least able to self assess accurately, with the poorest performers over-estimating their competence[4]” This statements does not make sense to me. An accurate and valid measurement of confidence and competence is important, but surely not because these two constructs are often poorly correlated, as claimed by the authors. Correlation between the two is not necessarily related to whether each is valid and accurately described, as confidence and competence will be expected to correlate poorly if they measure different constructs, even if both terms have been thoroughly validated. Here I think the authors should avoid citing poor correlation between the two as a reason to justify the need to better define confidence and competence. Also I think confidence is a subjective measure that may vary in different contexts even within the confine of measuring belief in clinical abilities, and so a uniformly accepted definition of confidence may not be possible. Questionnaire items: “The responses to questions “What makes a good and bad doctor and what makes you feel confident and competent?” “what makes you feel more confident or competent in dealing with children?” The wording of the question does not allow clear differentiation between the terms confidence and competence, they promotes confusion by suggesting that the two terms are actually interchangeable. It would have been more appropriate if each question was separated into two parts, one dealt with confidence and another competence. Discussion The authors alluded to the difficulties in measuring learning gain following educational intervention due to the difficulties in defining confidence versus competence. I suggest they include the Kirkpatrick model of evaluation in their discussion as it may provide an overall perspective in the measurement of learning gain.

REVIEWER	Alexandra Gilbert University of Leeds and St James' University Hospital, Leeds, UK
REVIEW RETURNED	17-Aug-2015

GENERAL COMMENTS	Review “Self evaluation of safety”: A more useful metric than confidence and competence? This is a very interesting article outlining the use of focus groups with junior doctors exploring their understanding of competence and confidence in clinical practice; in some groups focusing on a specific scenario of a febrile child. The research uses a meta-planning exercise methodology to examine these themes carrying out three focus groups over a two year period. Whilst the resulting paper has many interesting findings, the overall article would benefit from some re-structuring to highlight the key results. Specific comments: Abstract; page 2 1. ‘A perception of ‘safeness’ is as a’… remove word ‘as’ 2. There are a mixture of full stops and no full stops at the ends of each section. Strengths and Limitations; page 2 1. This is a first phases evaluation…’ I am not clear what this means. 2. ‘..which is able to visually describe…’ this sentence isn’t very clear. Perhaps it would be clearer to state that ‘participants develop and analyse the ideas discussed during the process’. Introduction 1. The introduction may benefit from re-structuring. Currently the key issues are only highlighted in the final paragraph whereas it would read more clearly if these concerns were injected throughout the background to guide the reader through. 2. Page 3, para 1: This section would benefit from describing specific examples of the use of self-assessment in clinical practice for example work based placed assessments. 3. The definitions of confidence and competence are not referenced. As the authors state these are often muddled it would be useful to state where these definitions are from as a starting point. 4. It would be helpful to describe how confidence and competence were mixed up in the Speechley et al example. 5. Page 3, para 2: It would be useful to give an example of a ‘descriptive qualifier’ of competence in this sentence. 6. Page 4, para 2: From reading the methods, the aims of the research appears wider than a focus only on the management of the feverish child as this was only used as an example in one of the focus groups. Is this correct? Methods 1. Page 4, para 3: Need to state if all FY academic trainees were invited to take part. 2. Perhaps it’s too much detail to explain how trainees split their time (with anatomy demostrations). 3. How trainees were approached (the final section of the methods) should follow on from who was invited to take part. 4. Add in the description of the metaplanning process ‘In summary

	the metaplanning approach…' after stating that this approach was used 'These doctors were asked…'. Then follow on with why it was used: 'This was used…' Results 1. Page 5 para 3: This first section of the results would be easier to follow if you re-structured to discuss the details about each focus group in turn e.g. Group 1, completing X date met on X date. X participants took part etc. 2. Appendix 1 only contains two sets of questions despite 3 groups taking part. Perhaps if you are not going to include all questions from all groups this appendix should just provide a few examples of the questions. 3. Table 1 (page 6-8) would benefit from reducing the results to the themes discussed, perhaps including a few key examples of the responses that illustrate the themes or these responses could be inserted as quotes in the text and reference between the table and the text. Currently having all of the responses makes the table difficult to read. Where the 2011 group were unable to develop themes for competence/confidence either you could extract them from the transcripts or leave without. This is a finding in itself that supports the argument that it's difficult to define competence and confidence and required you to change your methodology. 4. Figure 1 (page 9) is interesting but it would be helpful to have some further analysis of the figure in the text. 5. Page 9: Themes from discussion – in this section it is not clear which qualitative methodology was used to analyse the transcripts or if this is an extension of the meta-planning methodology. This needs stating. 6. Throughout the quotes some of the grammar makes it difficult to understand what the speaker is trying to describe: a. Page 10, quote 7: The first sentence is unclear b. Page 11, quote 6: The first sentence is unclear 7. Page 11-12; section on Alternatives to confidence and competence: There are too many quotes and little analysis in this section. This section would benefit from reducing the quotes and bringing in the interpretation of the model that is in the discussion – pages 14, para 3 to page 16. Currently having the interpretation in the discussion detracts from the bringing together the key messages of the study. Discussion 1. Page 13, para 2- check the quote '…whether they are "competent" (?to) give information…' 2. Pages 14, para 3 to page 16 – I would suggest moving the interpretation of the analysis to the results section as discussed above and restructuring the discussion following this change. Limitations 1. It may be worth re-stating that FY2s on the academic training programme were involved when describing them as a 'specific cohort' again in the limitations but perhaps defending why this was chosen – e.g. practical reasons such as ease of access, timing, availability to schedule meetings etc. 2. All trainees have to undergo similar WBPA and undergraduate curriculums so potentially all FYs should have a similar experience of reflective practice. 3. Page 17; para 4: Was informal feedback about the process provided to give you information on the 'time consuming' nature of the task? In this section mentioning the modification to the exercise

	for the 2012 group would be a good defence.

VERSION 1 – AUTHOR RESPONSE

Reviewer One

General Points: All three of these limitations have been addressed and noted in the manuscript. Two in the limitations section itself and one in the main discussion.

Introduction, paragraph 1
"An accurate and valid measure of confidence and competence is important because the least competent professionals are also the least able to self assess accurately, with the poorest performers over-estimating their competence[4]"
> This sentence has been removed to avoid confusion.

Questionnaire items:
"The responses to questions "What makes a good and bad doctor and what makes you feel confident and competent?"
"what makes you feel more confident or competent in dealing with children?"
> This was poorly explained in the initial manuscript and the separation of the terms has been highlighted.

Discussion
The authors alluded to the difficulties in measuring learning gain following educational intervention due to the difficulties in defining confidence versus competence. I suggest they include the Kirkpatrick model of evaluation in their discussion as it may provide an overall perspective in the measurement of learning gain.
> A clarification with reference has been made here.

Reviewer: 2

Reviewer 2
Specific comments:

Abstract; page 2
1. 'A perception of 'safeness' is as a'… remove word 'as'
2. There are a mixture of full stops and no full stops at the ends of each section.
> Amendments made

Strengths and Limitations; page 2
1. This is a first phases evaluation…' I am not clear what this means.
2. '..which is able to visually describe…' this sentence isn't very clear. Perhaps it would be clearer to state that 'participants develop and analyse the ideas discussed during the process'.

> Amendments made

Introduction
1. The introduction may benefit from re-structuring. Currently the key issues are only highlighted in the final paragraph whereas it would read more clearly if these concerns were injected throughout the background to guide the reader through.
> In conjunction with some of the points made by reviewer 1 alterations have been made.

2. Page 3, para 1: This section would benefit from describing specific examples of the use of self-assessment in clinical practice for example work based placed assessments.
> Amendment made

3. The definitions of confidence and competence are not referenced. As the authors state these are often muddled it would be useful to state where these definitions are from as a starting point.
> This is practically actually a challenge! but some references supplied

4. It would be helpful to describe how confidence and competence were mixed up in the Speechley et al example.
> This has been clarified

5. Page 3, para 2: It would be useful to give an example of a 'descriptive qualifier' of competence in this sentence.
> We are not sure what is being referred to here

6. Page 4, para 2: From reading the methods, the aims of the research appears wider than a focus only on the management of the feverish child as this was only used as an example in one of the focus groups. Is this correct?
> This has been clarified

Methods
1. Page 4, para 3: Need to state if all FY academic trainees were invited to take part.
> Amended
2. Perhaps it's too much detail to explain how trainees split their time (with anatomy demostrations)
> In the interests of word count and readability this has not been amended but can be done so if thought a vital piece of information

3. How trainees were approached (the final section of the methods) should follow on from who was invited to take part.
> amended

4. Add in the description of the metaplanning process 'In summary the metaplanning approach…' after stating that this approach was used 'These doctors were asked…'. Then follow on with why it was used: 'This was used…'
> Amended

Results
1. Page 5 para 3: This first section of the results would be easier to follow if you re-structured to discuss the details about each focus group in turn e.g. Group 1, completing X date met on X date. X participants took part etc.
> Amended

2. Appendix 1 only contains two sets of questions despite 3 groups taking part. Perhaps if you are not going to include all questions from all groups this appendix should just provide a few examples of the questions.
> Amended
3. Table 1 (page 6-8) would benefit from reducing the results to the themes discussed, perhaps including a few key examples of the responses that illustrate the themes or these responses could be inserted as quotes in the text and reference between the table and the text. Currently having all of the

responses makes the table difficult to read. Where the 2011 group were unable to develop themes for competence/confidence either you could extract them from the transcripts or leave without. This is a finding in itself that supports the argument that it's difficult to define competence and confidence and required you to change your methodology.
> Table reduced and additional information placed in an appendix

4. Figure 1 (page 9) is interesting but it would be helpful to have some further analysis of the figure in the text.
> This has been added to the discussion

5. Page 9: Themes from discussion – in this section it is not clear which qualitative methodology was used to analyse the transcripts or if this is an extension of the meta-planning methodology. This needs stating.
> Amended

6. Throughout the quotes some of the grammar makes it difficult to understand what the speaker is trying to describe:
a. Page 10, quote 7: The first sentence is unclear
b. Page 11, quote 6: The first sentence is unclear
7. Page 11-12; section on Alternatives to confidence and competence:
> amended

There are too many quotes and little analysis in this section. This section would benefit from reducing the quotes and bringing in the interpretation of the model that is in the discussion – pages 14, para 3 to page 16. Currently having the interpretation in the discussion detracts from the bringing together the key messages of the study.
> Some quotes have been removed to make this section easier to read.

Discussion
1. Page 13, para 2- check the quote '…whether they are "competent" (?to) give information…'
2. Pages 14, para 3 to page 16 – I would suggest moving the interpretation of the analysis to the results section as discussed above and restructuring the discussion following this change.
> This was considered but may reduce readability of the discussion. We hope the rationalisation of the results section means the sections are now clearer.

Limitations
1. It may be worth re-stating that FY2s on the academic training programme were involved when describing them as a 'specific cohort' again in the limitations but perhaps defending why this was chosen – e.g. practical reasons such as ease of access, timing, availability to schedule meetings etc.
> Amended

2. All trainees have to undergo similar WBPA and undergraduate curriculums so potentially all FYs should have a similar experience of reflective practice.
3. Page 17; para 4: Was informal feedback about the process provided to give you information on the 'time consuming' nature of the task? In this section mentioning the modification to the exercise for the 2012 group would be a good defence.
> Amended

VERSION 2 – REVIEW

REVIEWER	Nai Ming Lai School of Medicine, Taylor's University, Malaysia.
REVIEW RETURNED	13-Sep-2015

GENERAL COMMENTS	The manuscript is much-improved following the the author's revision to the comments. I see three of my four major points satisfactorily addressed in the text. One point that was not addressed the way the authors claimed they have is the following: “An accurate and valid measure of confidence and competence is important because the least competent professionals are also the least able to self assess accurately, with the poorest performers over-estimating their competence[4]” > This sentence has been removed to avoid confusion (authors' response) The sentence has not been removed as stated. However, the authors have elaborated in the subsequent texts well in their revision to justify the importance of validating the two constructs to render the sentence unnecessary. I will leave it to the authors to decide whether to remove or keep the sentence, and if they choose to keep the sentence, one suggestion is to move it to the end of the paragraph as the last sentence.

REVIEWER	Alexandra Gilbert St James' Institute of Oncology Bexley Wing, Level 3 St James' University Hospital Beckett Street Leeds West Yorkshire LS9 7TF
REVIEW RETURNED	24-Sep-2015

GENERAL COMMENTS	Thank you very much for the changes you have made, the article is much better. However, I was disappointed that you did not follow the recommendation to move the proposed matrices (table 3 and 4) to the results section as this describes a synthesis of the results not a discussion. This would leave space in the discussion to put the matrices into context. This move would help a clinician who is not familiar with qualitative methods to navigate the paper. The BMJ review article by Pope et al (2000) 'Analysing qualitative data' clarifies the importance of interpretation of the qualitative data as part of the analysis section. I also note that in a previous paper by the first author referenced in this article (20) the matrixes are proposed in the results section. The paper is otherwise very interesting and well research.This modification would make the research accessible to a much wider clinical audience.

VERSION 2 – AUTHOR RESPONSE

Reviewer One

I have uploaded both the new tracked changes and the previous tracked changes documents. It is clear that the sentence reviewer one mentions had been removed.

However it is similar to another sentence "Self-assessment is a complex, potentially learnt skill, requiring individuals to have insights into their own limitations and competencies(4)". Did reviewer one actually mean this sentence or was perhaps there a mistake in version control?

Reviewer Two

I had concerns that to include the matrix in the results would be to include some material that was clearly discussion rather than results. In order to address this issue Table 3 has been moved into the results section where as reviewer 2 suggests it does follow from the information presented. Table 4 remains in the discussion as a commentary on the applicability of the results in clinical practice.