Author’s response to reviews

Title: Impact of video feedback system on medical students' perception of their clinical performance assessment

Authors:

Bee Sung Kam (beesung@pusan.ac.kr)

So Jung Yune (cc139@pnu.edu)

Sang Yeoup Lee (saylee@pnu.edu)

Sun Ju Im (sunjuim11@hanmail.net)

Sun Yong Baek (sybaek@pnu.edu)

Version: 1 Date: 18 Jul 2018

Author’s response to reviews:

July 19, 2018

Editor-in-Chief

BMC Medical Education.

MEED-D-17-00837 Title: Impact of video feedback system on medical students’ perception of their clinical skills assessment
Dear Sir:

We are very grateful to you and the reviewers for the comments on the original version of our manuscript. This letter represents our response to the reviewer’s comments about our paper. We carefully considered all suggestions by the reviewer and submit a revised version of our manuscript. We highlighted with red color all changes made in the revised manuscript. Please find below a point by point response to editor’s and reviewer’s comments and clarify the important points of your main concerns. We believe these actions address the deficiencies and comments noted by the reviewer. We hope that you will be pleased with this revision and the revised manuscript will better meet the requirements of the ‘BMC Medical Education” for publication. We thank you again for the constructive review by the editor and reviewer.

Sincerely yours,

Sang Yeoup Lee M.D.

Family Medicine Clinic, Pusan National University Yangsan Hospital, Beomeo-ri, Mulgeum-eup, Yangsan, Gyeongsangnam-do, 626-770, South Korea

Phone: +82-55-360-1442, Fax: +82-51-510-8125, E-mail: saylee@pnu.edu

[Responses to the comments of Aaron William Bernard (Reviewer 1)]

Abstract:

Q1. Would remove "for a more effective feedback."

A1. Yes, we removed it.
Q2. I don't understand what "very few students were newly refused to agree with their scores after reviewing video feedback."

A2. We thank your very helpful comment. It means that there are rarely students who do not understand why such an evaluation result came out, and it seems that it is not properly described. From what you pointed out, we think this expression can confuse the reader. We described the meaning more clearly in the discussion section, instead we deleted it in the abstract.

Abstract and Introduction

Q3. You define the terms CSA, CPX, and OSCE in the methods in that you state the CPX and OSCE are components of a CSA. I think most people use all three terms interchangeably which makes this confusing.

A3. We appreciate your valid point of view and agree with your concerns. Therefore, to match the terms used in our recent study (Yune et al. Holistic rubric vs. analytic rubric for measuring clinical performance levels in medical students. BMC Med Educ. 2018 Jun 5;18(1):124), we have changed the term clinical skill assessment (CSA) to clinical performance assessment (CPA).

Q4. I don't quite understand line 15 "students do not always agree with their score." At first I thought, perhaps "students do not feel the score generated from the checklist is an accurate representation of their performance or skill." Having read the article further I am not sure if they disagree with how the checklist was filled out or if the checklist is a fair assessment of their skill.

A4. Thank you for your reasonable comments. As you pointed out, the present explanation seems to be somewhat confusing. We have amended to make it clearer as follows: “However, students do not always agree with their score.” □ “However, sometimes students do not feel that the score generated from the checklist is sufficient and properly assessed for their performance.”

Q5. The section discussing the different types of feedback is rambling. It could be good if made more concise.

A5. Yes. As your suggestion, we have amended those parts to make it more concise.
Q6. Line 55 "in the worse scenario a student might lose control." You are starting to get off track from the different types of feedback here and this could be removed.

A6. Yes, as your suggestion, we removed it.

Q7. Lines 16-28 or so are not clear to a reader that is not familiar with what you do. It is not clear what you used to do and how this new system is different.

A7. Yes. As your suggestion, we have amended those parts so that readers could better understand it.

Q8. Is the term "first grade medical students" common in other countries. In the US, the term is "first year medical students."

A8. Yes, we have changed the term "first grade medical students" to "first year medical students."

Q9. Can remove "because analyzed data retrospectively" The first sentence regarding IRB is enough.

A9. Yes, we removed “Because we analyzed data retrospectively and anonymously by assigning each student a distinct number, the institutional review board did not require informed consent from participants.”

Q10. Why was this not published since 2012? What was the delay? This makes me wonder if this paper has not made it through peer review at many journals.

A10. As you pointed out, we should have published this study as soon as possible, but due to various reasons, the completion of the paper was delayed and we have now submitted the paper. Nevertheless, we still think our ideas are unique in the field of medical education. Please look at this paper positively from that point.
Q11. The technical aspect of video compression is not of value to the reader

A11. Yes, as your pointing-out, we removed “This method of compression reduced the size of output data and therefore the time required for downloading the self-directed feedback was reduced.”

Methods and Results

Q12. It seems students reviewed both sets of videos without faculty involvement. I think this is a good idea but I'm not sure how you verify they watched either video. It could be better described.

A12. You may point out that we lack the description of the issue in the text. Unfortunately, we cannot force students to watch the videos and answer the questionnaire. However, after signing the informed consent form, we developed a program so that only students who responded to the first questionnaire were allowed to view their own video, followed by the recorded video of the best student, and then allowed the students to respond to the second questionnaire. We added this in the method section.

Q13. I'm not sure if student agreement with their score or comments is an appropriate outcome.

A13. We fully agree with what you have pointed out, and we also considered that. As is well known, today's medical education is oriented toward learner-centered education from teacher-centered education. In our thought, this is the same not only in education but also in feedback. We gave feedback via online-written comments even before implanting video feedback. This allows teachers to feel relieved that they have provided feedback, but does not look at whether students are receiving the feedback as well as the intention of teachers. Many teachers say: I said I taught him, I did not say he learned it. If students cannot do that, then the teacher did not do "teaching". (Whitman NA, Schwenk TL: The Physician as Teacher, USA: Whitman Associates, 1997.) Likewise, if the student does not agree to the feedback, then the feedback can be viewed as not feedback to the student. This study is one of a series of feedback studies. This research is to investigate students' perceptions of the professor's feedback. The results were as positive as we expected. In other words, as a result of video feedback, students who did not agree on their results before video feedback were reduced to 25.0% for CPX and 28.5% for OSCE after video feedback. Unexpectedly, a small number of students reported that they did not agree with the results of their evaluation after the video feedback was given. Further study such as a qualitative study may be required for some explanation of this findings. As you suggest, these points were further addressed in the introduction and discussion section.
Overall:

Q14. I think some time could be spent to have someone not familiar with the study read this and work with you to describe what you did to an outsider clearer.

A14. Thank you for your sincere and sympathetic comment. I understand what your intention is to comment. I made an overall revision so that readers could better understand this study.

Q15. I very much like the idea of this study but I question the outcome. Some people disagreed more after reviewing the video. I'm not sure agreement of score is equivalent to learning from the experience and if this is the correct marker.

A15. We fully understand your questions. Unexpectedly, a small number of students reported that they did not agree with the results of their evaluation after the video feedback was given. As you have doubts, we also wonder why the students who initially agreed to the results of their evaluation would not agree after seeing the video. Further study such as a qualitative study may be required for some explanation of this findings. However, most of the students agreed to the results of their evaluation after video feedback. We think this part is a new discovery of value. And when a student disagrees with a certain result, he/she would not accept any further direction or feedback either positive or negative. On the premise that accepting feedback from such a point of view, the feedback can influences learning process.

Q16. I think the unique thing you are doing that I would encourage you to follow up with in future scholarship relates to the use of an expert video that students could compare and contrast their performance too

A14. We thank your supportive comment. As you have recommended, we will continue to do further research.

Q17. Future work should look at verifying if students actually watched both sets of videos. Further description with a qualitative focus group study about what they learned maybe of value.

A17. Yes. We are also trying to extend this topic through further study of what you pointed out. Thank you for your advice.
Timothy J Wood (Reviewer 2):

Q1. The focus of this manuscript is solely on feedback but there is a small literature on the use of videos in OSCEs, which is relevant to this manuscript. I have listed some articles below. Note that this is not an exhaustive list but the author(s) should review some of them and include those that are relevant in their introduction.

A1. Thank you for your comment. We did not refer to the literature you presented because we looked at the recent literature, but we added it selectively according to your suggestion.

Cronin C, Cheang S, Hlynka D, Adair E, Roberts S. Videoconferencing can be used to assess neonatal resuscitation skills. Med Educ. 2001; 35(11): 1013-23. It seems to be inconsistent with the issue of this study because it is related to the remote online test.

Stone H. Angevine M. Sivertson S. A model for evaluating the history taking and physical examination skills of medical students. Med Teach. 1989;11(1):75-80. This is an appropriate reference. We quote it.

Harrison CJ, Molyneux SB, Blackwell S, Wass VJ. How we give personalised audio feedback after summative OSCEs. Med Teach. 2015;37(4):323-6. Although references to audio feedback have been provided, this is also an appropriate reference. We quote it.


Pinsky LE. Wipf JE. A picture is worth a thousand words: practical use of videotape in teaching. J Gen Intern Med. 2000;15(11):805-10. This paper is very relevant to the topic of this study and I would like to quote it.
Q2. I found the purpose of the study to be a bit weak. Rather than looking for improvements in performance given this feedback system, or study how the stakeholders use the information from the feedback system, or even compare different feedback systems, the purpose of the study was only to see if people agreed with the results. Perhaps the author(s) could better emphasize in the introduction why this is an important question to ask, and in the discussion provide some next steps they would implement to test out the benefits of this feedback system.

A2. We fully agree with what you have pointed out, and we also considered that. As is well known, today's medical education is oriented toward learner-centered education from teacher-centered education. In our thought, this is the same not only in education but also in feedback. We gave feedback via online-written comments even before implanting video feedback. This allows teachers to feel relieved that they has provided feedback, but does not look at whether students are receiving the feedback as well as the intention of teachers. Many teachers say: I said I taught him, I did not say he learned it. If students cannot do that, then the teacher did not do "teaching". (Whitman NA, Schwenk TL: The Physician as Teacher, USA: Whitman Associates, 1997.) Likewise, if the student does not agree to the feedback, then the feedback can be viewed as not feedback to the student. This study is one of a series of feedback studies. This research is to investigate students' perceptions of the professor's feedback. The results were as positive as we expected. In other words, as a result of video feedback, students who did not agree on their results before video feedback were reduced to 25.0% for CPX and 28.5% for OSCE after video feedback. Unexpectedly, a small number of students reported that they did not agree with the results of their evaluation after the video feedback was given. Further study such as a qualitative study may be required for some explanation of this findings. As you suggest, these points were further addressed in the introduction and discussion section.

Q3. On page 5, the information on how the questionnaire was developed is a sparse. What kind of relevant literature was reviewed, what kind of consensus was used. Why was a rating scale used given that all the questions in table 1 are really just yes/no questions. This does not need to have a large amount of text but a bit more detail is needed.

A3. We thank your very helpful comment. We added as follow: “The Conventional Delphi included questions being sent to a faculty focus group selected among members of clinical practice committee and, based on their responses, the Likert-scale questionnaire was developed and used to the study subjects.”
Q4. On page 6, there is mention of a pass/fail standard being applied. How the pass/fail standard determined and how many people failed? Similar to the point above, this does not need to have a large amount of text but a bit more detail is needed.

A4. Yes. We added some description to help readers understand a pass/fail standard in method section as follows: “A retake of an examination is permitted to students with the scores 1 SD below the mean. A Pass/Fail decision for CPA is based on the total scores 2 SD below the mean. Finally, on average, no more than three students have failed to pass this assessment.”

Q5. On page 7 there is mention of an ANOVA with the result displayed in figure 2. The problem is that I am not clear on what question was actually analyzed. Please clarify. Also on page 7 the author state that p-values between .05 and .10 will be labelled as marginally significant. Effects are either significant or not, there is no such thing as marginally significant. The author(s) may wish to use a more lenient p-value of .10 to declare significance but if so then they need to build a case for why they wish to do so. Personally, I would not do this but I suppose the authors could make a case for it. Finally, there are big issues around how to interpret p-values, and with 103 people relatively small differences will be statistically significant. I would suggest the authors report the effect sizes associated with all p-values to help the reader understand which questions are actually different and which are different just because there are so many people.

A5. Thank you for your advice. We deleted Figure 2 as the answer to Q8 below. We do not use marginal significance well. It is not used more and more in the biomedical science field. However, we think that the social science field such as education as below can be used occasionally to help readers to refer to this research result. And we added the effect sizes associated with all p-values to help the reader understand (see Method section and Table 1).


Q6. It was not clear to me until the discussion that the learners received a video of themselves along with the video of the highest performing learners. Up to that point, it seemed like learners only got the videos of the highest performer. Please ensure the methods are clear that this was the manipulation that was used.

A6. Yes. We will describe the manipulation more clearly in the method section.
Q7. For table 1, the order of the adjectives listed on page 6 and 7 are reverse of those listed in Table 1. Please correct. For the table perhaps add point on the scale corresponds to a 1, what corresponds to a 2 etc. Also specify that the table contains means and presumably standard deviations. I also wonder about the distributions. Based on table 2 and 3 it looks like hardly any 1’s and 2’s were chosen so these are highly skewed distributions ranging from 3-5. At this point, I start to wonder how much the scale is actually of value when distributions are so tight. Perhaps the authors could speak to the skewed distributions and why they provide useful information.

A7. Thank you for your comments. We described on page 7 the order of significance, marginal significance and agreement and perceptions of weaknesses presented in Table 1. The rest of it was reviewed again and the terms were uniformly modified in Table 1. And, the numbers listed in Tables 2 and 3 are not scale, but rather the number of students who did not agree with their results after each station was presented before and after video feedback. This is not a matter of the distribution.

Q8. Figure 2. I wonder if a figure is needed. Typically, production of figures is more costly than tables. That said, if a figure is used, please ensure the title, which labels the conditions by number, matches the actual figure, which labels conditions by letter. Also, what do the error bars represent? I also want effect sizes because looking at this figure, there is huge overlap in the error bars which leads me to think these are very small effects.

A8. We thank your comment. Figure 2 compares how much students agree with their own scores after the total CPA score report, online-written comment, and video feedback, which were selected from Table 1. However, as you pointed out, it’s redundant, so we will describe it in the text and delete Figure 2.