Author's response to reviews

Title: Educational Potential of a Virtual Patient System for Caring for Traumatized Patients in Primary Care

Authors:

Solvig Ekblad (Solvig.Ekblad@ki.se)
Richard F Mollica (RMOLLICA@PARTNERS.ORG)
Uno Fors (uno@dsv.su.se)
Ioannis Pantziaras (Ioannis.Pantziaras@ki.se)
James Lavelle (JLAVELLE@PARTNERS.ORG)

Version: 5 Date: 18 July 2013

Author's response to reviews: see over
Dear Editor of BMC Medical Education,

First we would like to take the opportunity to warmly congratulate you to your very interesting journal BMC Medical Education. It covers a broad area in the field of medical education. We are very pleased to realize that one of your interest areas encompasses simulation-based medical education – a field that we share. We hope therefore to contribute further to this area of medical education, by sending our revised version including Points by Point Response to each of the two reviewers, MS: 9168007469104745 “Educational Potential of a Virtual Patient System for Caring for Traumatized Patients in Primary Care” for your final consideration. Authors are: Solvig Ekblad, Richard F Mollica, Uno Fors, Ioannis Pantziaras and James Lavelle. To our knowledge, this is the first study examining these attributes in regard to a dedicated refugee trauma VP.

We would like to ensure you that this manuscript represents results from original work. The submitted material has not been published and it is not being considered for publication elsewhere. All authors have contributed significantly, have critically reviewed its content and have approved the final revised version submitted. Further, there are no conflicts of interest related to the research reported in this manuscript.

We are looking forward to your response and truthfully hope that this manuscript meets the standards of BMC Medical Education.

Yours sincerely,

Solvig Ekblad, PhD, Associate professor
Head of Cultural Medicine Unit
Corresponding author

Address to corresponding author:
Department LIME, Cultural Medicine Unit
Karolinska Institutet
Tomtebodavägen 18A, plan 3
SE-171 77 Stockholm, Sweden
E-mail: Solvig.Ekblad@ki.se
Point-by-point response

We thank the editor and the two reviewers for their valuable comments and suggestions! Below is our point-by-point response to the main revisions. Additionally, we have added/omitted single words in the text and it is easy to find those in track changes version. We hope that these revisions have improve the quality of our manuscript.

Best regards! Solvig Ekblad and co-authors

Our response is underlined

Pages and lines refer to the marked manuscript

Reviewer: 1

First of all, many thanks that you were intrigued by our paper and our ambition below is to improve the analysis and description of the educational tool by considering major, minor and discretionary revisions below.

MAJOR COMPULSORY REVISIONS:

1. More discussion of educational theory

Page 1, 6 lines from the bottom and page 2, first five lines from the top, we have revised and added the following text:

According to situated learning theory, students can acquire knowledge by being engaged in tasks that in an authentic way parallel real world activities [29]. Communication research has shown that highly engaging and challenging interactive media interfaces can promote deeper motivation and concentration [30]. Those factors have a major significant influence on the way we interact with others, e.g. authentic patient encounters and how we interpret verbal and non-verbal communication cues. Research in education has reported that emotional mediated experiences have a positive impact on cognitive learning outcomes [31, 32]. Thus, they lead to heightened involvement and decreased cognitive overload. Authenticity is critical to whether a virtual patient can be considered to be part of a situated learning endeavour, indicating that VPs may provide reliable, valid, and applicable representations of live patients [33].

See also below under Discussion, page 9, 5:e paragraph from the top.

2. Streamline the results and don’t overinterpret the very small dataset

We have excluded significance testing and only included Median and Ranges, in Table 1 and 2 and text in the Data analysis on page 5.
Page 5, the two paragraphs under Data analysis we have revised and added the following text:

We quantified all findings from the answers in the questionnaires in the pilot study using SPSS 19.0. Analysis of the quantitative data were primarily descriptive since the small N=11, does not allow for significance testing and will not allow for generalizability of gender and age. The analysis from the pre and post data included item-by-item measures and median values for measuring the average rating of the Likert scale questions included in the revised KI-VP-LEQ. Focus was on an evaluation of the realistic nature and usefulness of the VP-system when analyzing the participants’ answers. Evaluation of the face validity, defined as “the extent to which the examination resembles real life situations” [54] was analysed in the issue of acceptance of the degree of the realistic nature of the patient simulation in relation to the actual task as it has been used before including the Swedish pilot study [43].

The qualitative data included the answers from the participants during the 10 telephone interviews lasting between 15 and 30 minutes. The data was analyzed according to inductive content analysis based on Graneheim and Lundman’s model [55]. Anonymous citations will be used to exemplify the quantitative data. Sometimes the grammar has been changed to increase readability, but in a way that the content has not been changed.

Also in the Abstract we have changed the text under Result: The participants ranked the mental status examination more positively after the simulation exercise compared to before the simulation. Follow up interviews supported the results.

In the qualitative result part, we have streamlined the text and also made a number of language corrections in the citations, pp 6-8 as follows:

**General opinions and perceptions in the questionnaires after the learning experience**

Most of the participants were positive to the VP as being a great and helpful learning tool for exercising skills that appeared realistic and interactive. Some felt that it helped to expand the thinking in “Domains”.

- *It [VP] appeared real. When I started the program and was trying to figure out what to press, Katharina (VP) started to cough as a sign that she was waiting for me to start. I thought that was interesting.*

After the simulation exercise, the participants were more concrete in how to use VP in the future. They found VPs to be a great teaching tool and very valuable not only for medical but for any healthcare discipline. They considered that it was user-friendly and it could not only be used to assess patient's health but also patient’s resources. There were also critical comments

- *I liked it. It’s a great learning tool, many things can be done with it. Listening/watching a real person add a lot to the experience.*
- *I can’t think of any reasons why we should be using (this). There are better reasons to use it to improve our patient’s quality of regarding critique. Thu, its will improve our knowledge.*
Regarding the participants’ perceptions of virtual patients as compared to paper cases (Likert scale ranging from worse to better), seven of the 11 participants answered Median=5 (better).

- *Emotions, tone of voice, body language enhanced the experience [of VP]*.

The overall opinion about this learning method was positive: it was perceived as an easy and safe way to train relevant skills, good at various levels (beginners/medical students & experts/residents) and gave an opportunity to learn and refresh their interviewing skills of a full assessment and at the same time observe the patient’s reaction and response to their questions. Compared to hearing a lecture or reading relevant literature, VP was perceived as better. The video quality of the VP was mentioned to be good. Some recommended to have a summary in the end of the personal story “to point out a few general facts about how to approach others who have suffered trauma.”

- *When first working with people who are victims/survivors of trauma a clinician can be overwhelmed to the point of that they themselves being traumatized. Through this leaning method, the clinician can feel those moments within themselves, learn to deal with them, put the aside to a more appropriate place/time to talk about them etc. But initially those “Oh my God” moments maybe hard to deal with appropriately.*

The physical examination feature was evaluated as helpful by the participants but for some of them it was not considered as enough, as they wanted to pay more attention to gynaecological problems as well as reflexes and range of motions in arms, than they could. An additional limitation mentioned was that the participants could not physically touch the patient in order to i.e. auscultate her lungs or heart sounds, or feel her pulse. But they still felt it valuable to be able to watch the patient’s reaction:

- *...her frustration with waiting for the next question, her body language, her tone of voice all were key aspects to the assessment – certainly it added to the learning experience.*

Few of the participants felt that VP was an “ideal real patient”, and directed mostly to her psychological problems and reported this experience as follow:

- *Most people I see are further out from their refugee/trauma experiences than this patient, and less likely to have semi-acute medical issues like gonorrhoea.*

Others felt that after being trained with the VP, they became more open to ask people about their stories and to have full assessment of the patient’s status – physical, emotional, social and spiritual, which reinforces the complexity of patients.

It was common that the participants wanted to have the possibility to formulate their own questions in free text and experienced that scrolling among the pre-defined illness history questions in order to find the most suitable question was in some way odd. Some felt that the VP tool was somewhat artificial due to the limited options to ask questions in the order they wanted. They also reported that
the feedback of the VP and VA was not totally clear and needed to be technically improved.

All of the 11 participants gave constructive comments to improve the design of the VP system. Their suggestions included the possibility of asking their own questions without the need of scrolling among pre-defined questions, the technical improvement of feedback by the VP and Virtual Advisor and adding more extensive history details such as information about the patient’s daughter. Moreover, a more user-friendly interface that guides the user during the beginning and the ending of the case was suggested. One of participants summarized it like this:

- Several issues: (1) Traditional (folks) folder was empty; (2) Did not like read to scroll the questions; (3) unable to access background info – it did not open anywhere. Also the ending was confusing. I did put my assessment and submitted it, but did not get feedback. Then pressed wrong button and case was over. May be a warning “It would end your session, are you sure? Or would you like to hear feedback?

Feedback during the follow up interview regarding VP as authentic, acceptable, and ease to use

The participants recalled VP as a realistic and relevant virtual interface (“realistic responses to the questions”) with an interesting story that was very well done. Some remembered that the questions were very good and especially liked the four patient management domains. One participant hoped that this version could be rolled out for other languages and cultures, as it had great potential; the resource section needs a lot of work.

One of the participants remembered the VP as a useful tool especially for medical students but not for practicing doctors, as it was experienced as too artificial. It reminded some that most of their patients have some trauma in their lives and thought in a real patient would have had trouble dealing with all of that material.

- I think it was a very sad story, I feel compelled by all her losses. I hear similar stories from my patients and this story confirms my experience.
- I was moved by her. She reminded me of many of my patients.

A common answer from the participants (8 out of 10) was that VP could be used as an examination tool on many levels of competence, since it eliminates a lot of biases as everything which is used with this VP is the same.

- I think that my feeling is... it could eliminate a lot of biases because that is used with this VP is the same so the learner gets organized to think through the case.
- It gives you some distance to consider best questions and what directions should I go. The physical examination part one can’t do too much.

The participants mentioned that they would pay more attention to their interview skills after this training and that they would be more sensitive to the patient story and how emotional, psychosocial impact affects their health. They had increased awareness of the continuing effect of past trauma and were more willing to explore this concept deeper and not going too fast to traumatize the patient more. To get mastery with the four domains was also mentioned. One of the patients explored this in the following way:
- I am more aware of listening more attentively to patients.

The majority of the participants answered that they would recommend the VP to a colleague as a training tool in assessing a refugee trauma patient. Suggestions for improvement of the system were given by the participants also during the follow up interviews. They wanted it to be more user friendly, for instance giving the possibility to ask their own follow-up questions. The VA should be more easily accessible and more interactive and give more feedback about the participants’ interviewing skills. Further suggestions were that physical examination module needs to be expanded or should be skipped and some of the information was a little disjointed. It was recommended that the flow of questions should be increased and made more natural and with less structure. The participants would also have liked to be informed that it was not obligatory to ask all of the questions provided by the system.

- The patient and the answers were very real life. It was clear that a lot of work was put into it and it was very believable. However, the questions and the system were not easy to use with the click and submit methods.

3. Include in the discussion the implications of this tool now that you have demonstrated proof-of-concept. Can it be scaled up? What IT support will be needed?

Page 9, third paragraph from the bottom, page 10, the first three paragraphs from the top we have revised and added the following text:

However, the most common application of VPs is for learning and training clinical reasoning. Our VP educational model based PHC imply a cognitive constructivist pedagogy and a situated learning approach [29, 58] which may promote medical students’ patient-centered skills in interviewing for assessment, diagnosing and follow up. Good patient-centered interviewing skills have been connected with improved health outcomes and these strategies can according to Lein and Wills [59] “enhance effectiveness of patient care processes and outcomes while retaining efficiency of patient management” (page 215). Previous research has pointed out that VPs cases need to be realistic and preferably also based on real cases [61, 62]. The more authentic a virtual patient case is, it is more challenging [61], which may explain some of the reservations expressed by one of the participants who remembered the VP as a useful tool especially for medical students but not for practicing doctors, as it was experienced as too artificial. The reduction in post motivation scores in items 9, 11, 13 & 14 seems to show that the VP training was negative in some aspects, mainly because of the high expectations before the training procedure. This is called expertise reversal effect, i.e. the relative challenges in expert learners’ performance may happen “when there are overlaps between their well-learned and proceduralized knowledge structures and provided instructional guidance” [63, p. 333]. This evidence has significant implications for research in educational theory for adult, experienced learners as compared with novices, like medical students.
Practical implementation of VPs for clinical training
This study investigated how a small number of PCPs at a local clinic could make use of VPs for training their skills and even though none of them had used virtual patients before, they seemed to be able to run the RTSim system without any major problems. This indicates that VPs also might be used in clinical settings, where normally no teachers or facilitators are available. Therefore, we interpret our results that VPs might be used in for example CME in also other settings without too much need of support. This is in line with previous research [61], where CME has been indicated as one potential use of VPs.

The practical implementation of VPs for training has also been discussed in other studies, and indicated that it might be advantageous to use VPs in PBL settings or in other small groups as pairs of learners [64]. However, in clinical settings and/or for CME, individual use is probably the method of choice due to limitation in time and space for clinicians with heavy clinical burdens.

Limitations
One of the limitations in this study was the rather small N, which did not allow for generalizability of gender and age, which may have some impact on the PCPs’ perceptions of the utilization of the VP in PHC as suggested by the initial study [34].

The results were based on the subjective experiences of the participants without any external video observations, physiological recordings or other objective data [64] which could more deeply pay attention to the participants’ thoughts and attitudes to the VP.

There were some limitations due to technical issues which may have influenced the motivation of using the VP. Positive findings from the questionnaires included that the PCPs put more focus on mental status examination after the VP-based training. This is very positive since traumatized refugees often have mental impairment that needs to be identified, understood and taken into consideration by the PCP. Other non-significant, but still interesting positive findings included that the self-reported dimensions of clinical care showed positive trends in the post-test indicating that the PCPs tended to put more emphasis on root causes (social and spiritual) after the VP-training.

On the other hand, negative fluctuating datapoints in results came from the post-test in terms of that the participants disagreed more to use VP to help them to understand the mental health problems in their patients and to use VPs to help them to understand spiritual problems of their patients. These negative fluctuating datapoints might be connected with someone’s wish to apply open ended free text questions in order to build trust, which was not possible in the current version of the RTSim system. Such a comment was also given by one of the interviewed participants, who said that he lacked a feature to ask open ended questions to the Virtual Refugee. However, such VPs systems with free text dialogues are very expensive to develop, even with modern techniques.

MINOR ESSENTIAL REVISIONS:
4. Sort out the denominator issue in the presentation of the methods and results (it’s variably 11 or 10, though the methods section suggests it should be 10)
We have sorted out the numbers on page 4 under heading Participants, second paragraph from the bottom:
The 11 selected PCPs participated in the three 60 to 75 minute sessions led by the HPRT team leaders and received a US$ 225.00 dollar honorarium for each session. The participants were middle career practitioners (5 men and 6 women). All out of the 11 initial PCPs completed the study and returned completed questionnaires. One of the initial PCPs was unable to take part to the follow up interview due to illness.

(OBS not to be in the text: When the data from the interview was to be collected, the eleventh participant had died. However, due to integrity we will not write that in the text).

5. Review the one significance test used, it’s unclear what statistical test was used and the results seems unconvincing

On page 5 Data analysis, we have excluded significance testing, kindly also read above.

6. Don’t use the word “trend” for fluctuations in datapoints. The dataset is too small for a trend to emerge.

Page 10 under heading Limitation, second paragraph
We have excluded “trends” and re-phrased most of the texts indication significance or not..

DISCRETIONARY REVISIONS
In the discussion, consider: Is it best used in educational design for adult experienced learners or for medical students, and why? What of the argument that this is suited for standardised examinations rather than experiential learning?

Kindly read our comments on page 9, first and second paragraph from the bottom, page 10, the first six lines from the top we have revised and added the following text:

The reduction in post motivation scores in items 9, 11, 13 & 14 seems to show that the VP training was negative in some aspects, mainly because of the high expectations before the training procedure. This is called expertise reversal effect, i.e. the relative challenges in expert learners’ performance may happen “when there are overlaps between their well-learned and proceduralized knowledge structures and provided instructional guidance” [63, p. 333]. This evidence has significant implications for research in educational theory for adult, experienced learners as compared with novices, like medical students.

Practical implementation of VPs for clinical training
This study investigated how a small number of PCPs at a local clinic could make use of VPs for training their skills and even though none of them had used virtual patients before, they seemed to be able to run the RTsim system without any
major problems. This indicates that VPs also might be used in clinical settings, where normally no teachers or facilitators are available. Therefore, we interpret our results that VPs might be used in for example CME in also other settings without too much need of support. This is in line with previous research [61], where CME has been indicated as one potential use of VPs.

The practical implementation of VPs for training has also been discussed in other studies, and indicated that it might be advantageous to use VPs in PBL settings or in other small groups as pairs of learners [64]. However, in clinical settings and/or for CME, individual use is probably the method of choice due to limitation in time and space for clinicians with heavy clinical burdens.

INTRODUCTION: The paper would benefit from more engagement with educational theory. There is one sentence in the last paragraph under Discussion, which suggests the authors used a cognitive constructivist pedagogy and situated learning approach (unreferenced). Educational simulations are frequently presented as situated learning projects. Authenticity is critical to whether a virtual patient can be considered to be part of a situated learning endeavour. Arguably, authentic patient encounters are more challenging to simulate for experienced primary care practitioners than for medical student, which may go some way to explaining some of the reservations expressed by one of the respondents (Results, par 4). Could the authors expand their discussion of educational theory for adult, experienced learners as compared with novices, like medical students?

Kindly read our comments in Introduction and Discussion above under major compulsory revisions.

METHODS: There’s slippage in the denominator which the authors should clarity: it’s variously 10 or 11. If one of the participants was unable to continue because of illness, he/she should be included, unless participant 11 provided the post data but did not participate in the interview at a later stage. I had to return to Ref 34 to understand the evaluation methods. That reference explicitly states that no significance tests were used (correctly) because of the small numbers. I am unclear what significance test was used to establish a significant difference between the 10, or 11, participants in before and after in mental health assessment; but it seems unconvincing, given the small numbers.

We have sorted out the numbers on page 4 under heading Participants, second paragraph from the bottom:
The 11 selected PCPs participated in the three 60 to 75 minute sessions led by the HPRT team leaders and received a US$ 225.00 dollar honorarium for each session. The participants were middle career practitioners (5 men and 6 women). All out of the 11 initial PCPs completed the study and returned completed questionnaires. One of the initial PCPs was unable to take part to the follow up interview due to illness.

(OBS! When the data from the interview was to be collected, the eleventh participant had died. However, due to integrity, we do not want to write that in the text.)
RESULTS: To my mind the presentation of the results is over-extended. This section would read better if it were compressed and arranged around proof-of-concept issues, that is focusing on it’s the authenticity of the educational experience, acceptability, penetration (as evidenced by recall and recognition of key foci) and ease of use. Some of the quotes are hard to interpret, particularly the quote in the last sentence par 4 of the Results section. Other quotes are grammatically odd and require the reader to decode them. The Likert scale data mean little given the low numbers and are accorded too much interpretive weight. The use of the word “trend” in the Limitations section is misleading. There can’t be before-after trends with such a small dataset; what we’re looking at is fluctuating datapoints.

Page 5, the third and fourth paragraph from the top, we have revised and added the following text:

We quantified all findings from the answers in the questionnaires in the pilot study using SPSS 19.0. Analysis of the quantitative data were primarily descriptive since the small N=11, does not allow for significance testing and will not allow for generalizability of gender and age. The analysis from the pre and post data included item-by item measures and median values for measuring the average rating of the Likert scale questions included in the revised KI-VP-LEQ. Focus was on an evaluation of the realistic nature and usefulness of the VP-system when analyzing the participants’ answers. Evaluation of the face validity, defined as “the extent to which the examination resembles real life situations” [54] was analysed in the issue of acceptance of the degree of the realistic nature of the patient simulation in relation to the actual task as it has been used before including the Swedish pilot study [43].

The qualitative data included the answers from the participants during the 10 telephone interviews lasting between 15 and 30 minutes. The data was analyzed according to inductive content analysis based on Graneheim and Lundman’s model [55]. Anonymous citations will be used to exemplify the quantitative data. Sometimes the grammar has been changed to increase readability, but in a way that the content has not been changed.

Page 10 under heading Limitation, second paragraph
We have excluded “trends” and included according to your recommendation “fluctuating datapoints.

DISCUSSION: Some of the respondents suggest that the VP case may be useful for examinations. The authors, on the other hand, intend this to be a discovery learning exercise for experienced learners. Is it possible that this method may have a role as a standardised testing exercise, rather than a learning exercise; and if so, does this change the design approach? While the addition of trauma to the DMM seems like a good idea, it isn’t related to the evidence in this paper. On the other hand, a case for this may emerge from the forthcoming RCT.

Page 9, 13 lines from the bottom and to the bottom of the page, page 10, the first six lines from the top we have revised and added text, kindly read above
As this is a proof concept, how easy would it be to sale this up? What IT resourcing is likely to be needed? Can this be done with no input from an actual facilitator, and is it possible that this method could be used by two learners working together?

We have inserted a section on this on page 9-10 as follows:

**Practical implementation of VPs for clinical training**

This study investigated how a small number of PCPs at a local clinic could make use of VPs for training their skills and even though none of them had used virtual patients before, they seemed to be able to run the RTSim system without any major problems. This indicates that VPs also might be used in clinical settings, where normally no teachers or facilitators are available. Therefore, we interpret our results that VPs might be used in for example CME in also other settings without too much need of support. This is in line with previous research [63], where CME has been indicated as one potential use of VPs.

The practical implementation of VPs for training has also been discussed in other studies, and indicated that it might be advantageous to use VPs in PBL settings or in other small groups as pairs of learners [64]. However, in clinical settings and/or for CME, individual use is probably the method of choice due to limitation in time and space for clinicians with heavy clinical burdens.

**Level of interest:** An article whose findings are important to those with closely related research interest.

**Quality of written English:** Needs some language corrections before being published.

We sent the manuscript to a translator office to check the English before we sent in the manuscript. We have revised citations, as we understand that was your comment. Two of the co-authors have English as mother tongue.

**Statistical review:** No, the manuscript does not need to be seen by a statistician.

**Declaration of competing interests:** I declare that I have no competing interests.
Reviewer 2
We also would like to thank for constructive comments and our ambition is to increase the quality of our manuscript.

Discretionary Revisions
1.- The educational value of interactive computer simulations with virtual patients is adequately mentioned and some relevant studies are cited, although some others are missing.

We have added some of them and you may find them as the following numbers in the reference list: 39, 40, 41, 42, 64.

Additional suggested references are:
2. The paper is focused on subjective opinion data. It is interesting to see that this opinion is predominantly positive, although the objective assessment of the learned skills should also be reported.

Under the heading Limitations, page 10, second paragraph we have revised and added text:

The results were based on the subjective experiences of the participants without any external video observations, physiological recordings or other objective data which could more deeply pay attention to the participants’ thoughts and attitudes to the VP.

Major Compulsory Revisions
1. Even though the pre and post-test questionnaires are described in detail in a previous publication, a wider description than the one provided is also needed in this paper.

On page 5, heading 1 Questionnaires under Data collection the first ten lines from the top and text until the heading 2) A follow-up interview we have revised and added the following text:

The pre and post questionnaires included survey questions on the PCP’s IT proficiency, clinical worldview, and current motivation to use VPs in the learning exercise. A similar questionnaire set was also used in the Swedish pilot study with good results, making us believe that the questions had acceptable validity. The pre-test version included two parts. The first (“Overview of Clinical Worldview”) with the aim to examine the participants’ self-reported important issues during a “real life” medical examination, on a scale from 1 (no emphasis) to 5 (full emphasis). It consists of 10 items, divided into two parts; regarding the level of emphasis the clinician usually places on (a) data (chief complaint, history of present illness, physical exam, mental status exam, laboratory tests, and traditional healing exam) collected during the medical examination and (b) root causes of the disease (biological, psychological, social/economic and spiritual). The second part of the pre-test questionnaire had the aim to examine the current motivation to use the VP before the simulation exercise on a 4-point Likert scale (1= highly disagree, 4= highly agree) with 17 questions (eg. “I am motivated to use VP as it helps to improve interdisciplinary communication”, “I believe that VP will help me to provide better care to my traumatized patients from any cultural background”).

2. The specification of the statistical test applied to the results included in table 1 is required.

Page 5, seven lines from the bottom, page 6, the first eight lines from the top we have revised and added the following text:
We quantified all findings from the answers in the questionnaires in the pilot study using SPSS 19.0. Analysis of the quantitative data were primarily descriptive since the small N=11, does not allow for significance testing and will not allow for generalizability of gender and age. The analysis from the pre and post data included item-by-item measures and median values for measuring the average rating of the Likert scale questions included in the revised KI-VP-LEQ. Focus was on an evaluation of the realistic nature and usefulness of the VP-system when analyzing the participants’ answers. Evaluation of the face validity, defined as “the extent to which the examination resembles real life situations” [54] was analysed in the issue of acceptance of the degree of the realistic nature of the patient simulation in relation to the actual task as it has been used before including the Swedish pilot study [43].

The qualitative data included the answers from the participants during the 10 telephone interviews lasting between 15 and 30 minutes. The data was analyzed according to inductive content analysis based on Graneheim and Lundman’s model [55]. Anonymous citations will be used to exemplify the quantitative data. Sometimes the grammar has been changed to increase readability, but in a way that the content has not been changed.

In the qualitative result part, we have streamlined the text and also language corrections in the citations, pp 6-8

Also in the Abstract we have changed the text under Result: The participants ranked the mental status examination more positively after the simulation exercise compared to before the simulation. Follow up interviews supported the results.

3. **Also, the specification of the statistical test applied to the results included in table 2 is required**

See above, we have excluded statistical test due to the small numbers of participants.

4. Even though there were no significant changes between the pre and post-tests of self-reported motivation to use VP, a deeper discussion of these results is required. In particular, the reduction in post motivation scores in items 9, 11 and 13 seems to show that the VP training was negative in some aspects, mainly because of the high expectatives before the training procedure.

On page 8 under heading Discussion, first paragraph7:e line from the heading we have revised and added the following text:

The participants ranked the mental status examination more positive after the simulation exercise compared to before the simulation. Only minor changes was found in datapoints between the pre- and post-tests of self-reported motivation to use VP as an educational tool. Although the general motivation as ranked during the pre-test was high, the participants exhibited high expectations and a positive attitude towards the VP system. Some items in post-test scored however lower in
the post-tests (including items 9, 11 and 13) which can be interpreted as that the participants acquired a more realistic view on the current version of our VP having its limitations and areas that need improvement.

We further expanded our discussion on page 9, third, fourth, fifth and sixth paragraph from the top and page 10 first paragraph from the top:
Most of the respondents suggested that the VP case may also be useful for exams. This is positive and something that is used as a routine nowadays at both medical schools and as board exams [cf. 37, 38]. This also indicated that the PCPs in our study believe the RTSim system is capable of visualizing cases that resemble real patient cases.
However, the most common application of VPs is for learning and training clinical reasoning. Our VP educational model based PHC imply a cognitive constructivist pedagogy and a situated learning approach [29, 58] which may promote medical students’ patient-centered skills in interviewing for assessment, diagnosing and follow up. Good patient-centered interviewing skills have been connected with improved health outcomes and these strategies can according to Lein and Wills [59] “enhance effectiveness of patient care processes and outcomes while retaining efficiency of patient management” (page 215). Previous research has pointed out that VPs cases need to be realistic and preferably also based on real cases [61, 62]. The more authentic a virtual patient case is, it is more challenging [61], which may explain some of the reservations expressed by one of the participants who remembered the VP as a useful tool especially for medical students but not for practicing doctors, as it was experienced as too artificial.
The reduction in post motivation scores in items 9, 11, 13 & 14 seems to show that the VP training was negative in some aspects, mainly because of the high expectations before the training procedure. This is called expertise reversal effect, i.e. the relative challenges in expert learners’ performance may happen “when there are overlaps between their well-learned and proceduralized knowledge structures and provided instructional guidance” [63, p. 333]. This evidence has significant implications for research in educational theory for adult, experienced learners as compared with novices, like medical students.

Practical implementation of VPs for clinical training
This study investigated how a small number of PCPs at a local clinic could make use of VPs for training their skills and even though none of them had used virtual patients before, they seemed to be able to run the RTSim system without any major problems. This indicates that VPs also might be used in clinical settings, where normally no teachers or facilitators are available. Therefore, we interpret our results that VPs might be used in for example CME in also other settings without too much need of support. This is in line with previous research [61], where CME has been indicated as one potential use of VPs.
The practical implementation of VPs for training has also been discussed in other studies, and indicated that it might be advantageous to use VPs in PBL settings or in other small groups as pairs of learners [64]. However, in clinical settings and/or for CME, individual use is probably the method of choice due to limitation in time and space for clinicians with heavy clinical burdens.

5. Only four references are found in the discussion section, and only one of those
references is directly related with the field of virtual patients (and it is a previous study of the same authors). A wider and deeper elaboration of this section is needed, in order to compare the results obtained in this study with other similar studies that can be found in the scientific literature (some of them have been listed in the “Discretionary Revisions” section of this review).

We have added references, eg. 29, 58, 60, 61, 63, 64. In the Introduction we have added 39-42

Level of interest: An article whose findings are important to those with closely related research interests

Quality of written English: Acceptable

Statistical review: No, the manuscript does not need to be seen by a statistician.

Declaration of competing interests: I declare that I have no competing interests