Author’s response to reviews

Title: High-fidelity is not superior to low-fidelity simulation but leads to overconfidence in medical students

Authors:

Christina Massoth (christina.massoth@t-online.de)
Hannah Röder (hroeder92@googlemail.com)
Hendrik Ohlenburg (ohlenburg@anit.uni-muenster.de)
Michael Hessler (michael.hessler@ukmuenster.de)
Alexander Zarbock (alexander.zarbock@ukmuenster.de)
Daniel Pöpping (poppind@uni-muenster.de)
Manuel Wenk (manuelwenk@uni-muenster.de)

Version: 1 Date: 06 Jan 2019

Author’s response to reviews:

Editor Comments:

Dear Christina Massoth,

please use subheadings for to structure the methods section. These subheadings should mirror subheadings in in the results section. Whenever possible, please try to avoid one-or-two-sentence paragraphs, as this impedes a good flow of reading. Please include critical comments regarding the assessment of Knowledge to assess clinical Performance (miller pyramid) and the literature that doubts a correlation between self-assessment and clinical performance.

Answer: Thank you for your helpful advice to improve our manuscript. We revised the manuscript with regard to the formal specifications and added the following sections, addressing your content-related comments:

Lines 177-185:
Using the framework of the classical Miller pyramid of clinical competence assessment with its ascending levels from knowledge at the base to action at the top (knows-> knows how->shows how->does),(27) a relationship between specific forms of simulation training and correspondingly achievable levels of competency has been described. However, even if a high degree of realism favours the acquisition of the highest level of the pyramid (“does”-action), it requires for the condition of more qualified trainees proportionally to the degree of realism, to surpass the level of “shows how”. (28) Conversely, this may suggest that trainees at a lower educational level, as medical students in particular, are more likely to benefit from lower degrees of simulation training, supporting the results from this study.

Lines 216-221:

The ability of medical students to self-assess is known to be limited even though it seems to improve in accuracy in later years of medical school (35) However, inaccuracy of self-assessment is not only a problem of medical students: irrespective of the level of training the relationship between self-assessment and external assessment was found to be weak (36) Further, self-rated assessment is not only an inappropriate predictor of actual performance (37), those who self-assess more inaccurately are also more likely to perform weakly. (36,38,39)

Reviewer 1 Comments to the Author:

Very good study. Well written and commendable. Kindly find my comments/questions below:

Methods:

1. Why 4th year students alone? Could their year of study affect assimilation/knowledge/skills? This can be added to the study limitations if the year of study could affect the results.

Answer: We would like to thank the Reviewer for his kind words regarding our manuscript.

We agree with the Reviewer: the educational and training levels of 4th year medical students in particular may have affected the results of our study- we now address this point in the discussion (lines 180-185, lines 216-217) and conclusion (lines 256-258) and added it as a limitation (lines 240-243).
2. Why misinform the participants about the study purpose? Was it really necessary?

Answer: We considered students’ knowledge of the true purpose as a potential bias and were concerned about a possible distortion of results. Therefore, we anticipated to achieve a higher quality of results by blinding the participants with regard to the intervention.

3. Confidence was self-rated and those exposed to HF may naturally rate themselves higher. It may be better to use more objective tools to assess confidence in future studies.

Answer: Thank you for this important hint. We completed this point to our limitations.

4. Results:

Ventilation without equipment (29 vs 41%) and placement of venous cannula (26 vs 39%) were much better in the HF group.

Could these have been significant if analyzed differently? Was there nothing the HF group could do better?

Answer: We used chi-squared test for the analysis of these items and consider it as the test of choice for this question. We do not consider it useful to perform further test methods, being potentially less appropriate until a statistical significance is found. However, the HF-group performed better in several other items, but none of the differences was statistically significant.

5. Conclusion:

Although your study showed non-superiority of HF, science is dynamic and we must continuously seek improvements.

Moving ahead, although not part of your study, are there areas of HF simulation you identified that if improved will lead to better performance? Could these be included as recommendations?
Answer:

Thank you for this interesting comment. Though, as this was indeed no part of our study we did not reflect this point in detail in the revised version, but considered it by discussing an adjustment of the degree of fidelity to the educational level, suggesting a more differentiated use of HF, as not “one size fits all”, but trainees at a higher level of education may be more likely to benefit from HF simulation. (Line 179-185).

Reviewer 2 Comments to the Author:

1. Methods:

A cohort of undergraduate students, usually follows a Bell curve (in terms of performance), into the majority being average and lesser high achieving and borderline (low achieving) students. Is it known whether if in this study the majority of high achieving students were randomly allocated in the LF simulation group? What was done to ensure that all group were homogeneous in performance.

Answer:

We agree with the reviewer with regard to the assumption of a normal distribution. Also, combinations of students with similar performances in subgroups of the cohort are expectable, therefore students were allocated separately to the respective intervention group (HF or LF) using a randomization sequence with the method of permuted blocks. No further matching was performed, as we considered the process of randomization as sufficient to ensure adequate homogenization.

2. Also were the students in the HF group given adequate time to familiarise themselves with the equipment, as the initial learning curve is usually higher with HF simulators as compared with LF simulators which can account for some delays and decrease in performance.

Answer:

We agree with the reviewer that the use of high-fidelity simulators in complex scenarios in general have initially a smaller learning curve. ALS training, however, includes the realization of a relatively simple algorithm within a clear situation. In this regard, we do not believe that the
use of a more complex mannequin and environment requires for a considerable longer period of introduction, therefore both groups received similar time periods for familiarizing.

3. Results:

If the cohort in the study consists of millennial learners then there is a natural gravitation towards technology (Keengwe J, Georgina D; 2013), hence their preference for choosing HF in the cohort.

Answer: Thank you for this remark. We added a corresponding reference to the discussion section in the revised version of the manuscript.