**Author’s response to reviews**

**Title:** The Collaborative Learning Development Exercise (CLeD-EX): an educational instrument to promote key collaborative learning behaviours in medical students

**Authors:**

Maha Pervaz Iqbal (m.pervaziqbal@unsw.edu.au)
Gary Velan (g.velan@unsw.edu.au)
Anthony O’ Sullivan O’ Sullivan (A.OSullivan@unsw.edu.au)
Chinthaka Balasooriya (c.balasooriya@unsw.edu.au)

**Version:** 1  **Date:** 29 Jan 2020

**Author’s response to reviews:**

Dear Editor,

We are pleased to resubmit for publication the revised version of MEED-D-19-01211 “The Collaborative Learning Development Exercise (CLeD-EX): an educational instrument to promote key collaborative learning behaviours in medical students.” We appreciate the constructive comments, suggestions and queries of both the reviewers. We have addressed each of their comments as outlined below.

**Reviewer reports:**

Umatul Khoiriyah, Ph.D (Reviewer 1) comments and suggestions: The manuscript describes the development of a collaborative learning instrument namely collaborative learning development exercise (CleD-Ex). The manuscript has been written systematically, however; some remarks are identified to improve the paper.

Authors response: The authors would like to thank the reviewer for the positive comments and providing suggestions to improve the manuscript.

Umatul Khoiriyah, Ph.D (Reviewer 1) comments and suggestions: The main objective of the research and the research question were not stated clearly in the abstract as well as in the main text. Was the aim of the research to evaluate the validity of the CleD-Ex or were there other objectives?
Authors response: The authors thank the reviewer for providing this suggestion. We have added a sentence in the abstract to clarify the aim of this research reported in this manuscript, as follows: “In this article we report on the effectiveness, feasibility and educational impact of the CLeD-EX.” (revised line number 43-44, highlighted in yellow). This research was part of the wider project on collaborative learning in medical education and in this paper, we report on CLeD-EX instrument which was developed and its effectiveness, feasibility and educational impact. We note that we have clearly stated this in the background section of the manuscript (revised line 120-121, highlighted in yellow).

Umatul Khoiriyah, Ph.D (Reviewer 1) comments and suggestions: Most of the content of the paper showed the process of development and validation of CleD-Ex, however; the validity framework applied in this research was not clear. It would be good if the authors could describe the specific framework for instances Kane's validity framework or other current validity framework recommended in medical education research (i.e. Cook & Beckman, 2006). By using a specific framework, the research would be more rigorous and systematic.

Authors response: The authors thank the reviewer for this suggestion. However, as clarified in the aim of the research, reported in this manuscript, we have focused on the effectiveness, feasibility of use and educational impact of the CLeD-EX instrument rather than focus on the validity of the CLeD-EX. Therefore, we have not used a validity framework to report the findings.

Umatul Khoiriyah, Ph.D (Reviewer 1) comments and suggestions: In my opinion, the concept of collaborative learning described in the background needs to be more sharpen especially in terms of its connection with the other learning concepts such as critical thinking skill and self-directed learning. It is important since the authors measured the impact of the CleD-Ex by evaluating critical thinking and self-directed learning skill.

Authors response: As per the reviewer suggestion we have expanded the discussion to link the concept of collaborative learning with other learning concepts like self-directed learning and critical thinking. We have updated and added the following: “Therefore, during collaborative learning students have the opportunity to develop skills in managing their learning (i.e. self-directed learning) while interacting with peers to discuss and debate concepts which will promote higher order cognitive reasoning.” (revised line 78-80, highlighted in yellow).

“Moreover, self-evaluation can be valuable in group work assessment especially when used for the purpose of formative assessment (15). Self-evaluation also allows students the opportunity to identify their learning needs while taking increased responsibility for their learning, i.e. developing their skills in self-directed learning (16).” (revised line 101-103, highlighted in yellow).
Umatul Khoiriyah, Ph.D (Reviewer 1) comments and suggestions: In the methods section

The authors mentioned the modified Delphi technique in this section. Even though it has been published in previous research, it would be good if the author could describe shortly regarding the expert contributed to this Delphi process.

Authors response: As per the reviewer recommendation, the following has been added about the expert contribution in the Delphi study: “All the experts involved in the two-round Delphi (round 1 n= 54; round 2 n= 23) had an academic role in universities across Australia and New Zealand and were involved in small-group teaching that included design and/or facilitation of collaborative learning.” (Revised Line 131-134 highlighted in Yellow).

Umatul Khoiriyah, Ph.D (Reviewer 1) comments and suggestions: It would be useful for the reader if the authors could describe all the research process in figure 1. I am not quite clear how many times students filled the CleD-Ex, how long the process from stage 1 to stage 3 was conducted, the information whether the control group receive feedback or not and other information in term of the research method

Authors response: As per the reviewer suggestion, we have updated Figure 1 (revised line 191-197, highlighted in yellow). We have included the time durations and have added notes on students completed the CLeD-EX once during the course. Students in both the intervention and control group had the opportunity to complete CLeD-EX which involved receiving tutor feedback.

Umatul Khoiriyah, Ph.D (Reviewer 1) comments and suggestions: Result section

The Cronbach alpha of students' evaluation was only 0.451. It seems that this CleD-Ex does not enough validity evidence for its construct. Please clarify for this result.

It would also be good if the authors show the reliability of each item in CleD-Ex.

Authors response: We would like to thank the reviewer for this comment and suggestions. As per the reviewer comment and suggestion, we have included further clarification on the Cronbach alpha of 0.451 in the student self-evaluation scale in the discussion section: “The internal consistency of the CLeD-EX was greater in the tutor evaluation scale. The low Cronbach alpha in the student evaluation scale might be related in part to low sample size in the field-test.” (Revised line 388-392, highlighted in yellow). Also, as per the review suggestion we have reported on the reliability of each item in CLeD-EX. Therefore, we have updated information in both the results and discussion (revised line 248-253 and 388-392, highlighted in yellow).
Umatul Khoiriyah, Ph.D (Reviewer 1) comments and suggestions: Discussion section

The analysis regarding the discrepancy between tutor and self-evaluation needs to be more comprehensive. The authors should consider the other factors contributing to this condition.

The analysis regarding the connection of collaborative learning and critical thinking and self-directed learning also needs to be more elaborated with the theory.

Authors response: We thank the reviewer for these suggestions to improve the discussion on the discrepancy between tutor and self-evaluation and the connection of collaborative learning with critical thinking and self-directed learning, reported in the manuscript. Therefore, as per the reviewer comments we have revised and added the following information in the discussion: “The literature also reports poor correlation between self-evaluation and tutor evaluation in a small group learning context (17, 42).” (revised line 396-397, Highlighted in yellow)

“There is a tendency for weaker students to overestimate their work and for better students to underestimate their performance (43-45). Therefore, students’ self-evaluation may vary in accuracy. Self-assessment may thus be of limited value, if our aim is accuracy. However, in the context of developing essential professional skills of self-evaluation, self-assessment could play a significant role. Thus, self-assessment could be valuable in itself, but could play an enhanced role when paired with a peer or tutor assessment.” (Revised line 400-405, Highlighted in yellow).

“Learning-oriented interactions in small group learning environments are reported to be constructive, collaborative and self-regulatory (46). In the evaluation of CLeD-EX as an educational instrument, we explored the impact of the instrument on students’ self-directed learning, their ability to think critically and to interact within the group.” (Revised 411-414, Highlighted in yellow)

The authors mentioned that this tool designed for the busy educational context. I noted in the analysis of utility and feasibility part (page 10) tutors provided feedback around 2-15 minutes for each student. Does it mean that the feedback process could take time around 100 minutes for a group consisting of 10 students? Please clarify this issue since it is one of the reasons for designing this new tool.

Authors Response: We would like to clarify on the feasibility of CLeD-EX as an educational instrument which can be used in busy educational contexts. In addition to the feasibility of CLeD-EX in terms of the tutor time taken for providing feedback, CLeD-EX provided a framework for tutors to use for giving student feedback, in that they focused in promoting the key collaborative learning behaviours in students. In the discussion section, we note that we have discussed about this: “Learners often report receiving insufficient feedback from tutors and whatever feedback is received is often not helpful, timely or constructive (41). Important reasons for these findings could be time constraints faced by teachers or lack of experience or structure in providing feedback to their students……... Therefore, CLeD-EX aimed to address these issues by developing a process that would minimise the extra time load imposed on tutors, while providing a framework to give students feedback on these aspects.” (Revised line 376-384). In addition to
this, we have added further clarification about the feasibility of using CLeD-EX as an educational instrument: “CLeD-EX in small group learning could provide the framework for tutors to engage with the students who need feedback the most; that is the 2-15 minutes time range implies that two minutes can be sufficient in many cases, but tutors can choose to spend additional time with one or two students who need the extra attention.” (revised line: 379-382, highlighted in yellow)

Adam Wilson (Reviewer 2) comments and suggestions: In the manuscript the authors attempted to evaluate an exercise which fosters the development of collaborative learning competencies in medical students. Overall the manuscript was well written and easy to follow. My recommendations and comments are outlined below.

Authors response: The authors would like to thank the reviewer for the positive comments and providing suggestions to improve the manuscript.

Adam Wilson (Reviewer 2) comments and suggestions:

Introduction

First paragraph - throughout consider using the term "collaboration" skills instead of "collaborative" skills.

Authors response: As per the reviewer suggestion, this has been amended in the first paragraph of the background section (revised line 70, highlighted).

Adam Wilson (Reviewer 2) comments and suggestions: First paragraph, third sentence, after the comma - consider saying "elevating the importance of collaborative learning." (i.e., remove "therefore students require the capability to collaboratively learn together.")

Authors response: As per the reviewer suggestion, this has been amended in the first paragraph of the background section (revised line 73, highlighted).

Adam Wilson (Reviewer 2) comments and suggestions: Please write out CLeD-EX the first time it is used in the introduction.

Authors response: As per the reviewer comment, this has been amended in background section (revised line 117, highlighted).
Adam Wilson (Reviewer 2) comments and suggestions: The research question needs to be clearly stated.

How do you measure "impact", as this term can take-on a variety of different meanings?

Authors response: We would like to thank the review for the suggestion. We note that we have clearly stated the aim of the research which has been reported in the manuscript: “In this manuscript we report on the effectiveness, feasibility and educational impact of the CLeD-EX.” We have also clarified that in this manuscript we aim to explore the educational impact of CLeD-EX. (revised line 120-121).

Adam Wilson (Reviewer 2) comments and suggestions: Methods

General comment: The intervention in this study needs to be clearly defined. Was the intervention the assessment process or the session exercise? At times you switch back and forth between calling CLeD-EX an instrument versus calling it a process. The two are very different, as the validity of a process is measured differently than the validity of an instrument. Please clarify.

Authors response: We thank the reviewer for this important comment. We clarify that CLeD-EX is an assessment instrument which is introduced in this manuscript. Therefore as per the reviewer comment, we have revised all statements which describe CLeD-EX as a process (revised line 141, 433 and 456)

Adam Wilson (Reviewer 2) comments and suggestions: I applaud the authors for using a Delphi approach to discern the key behaviors to be assessed. However, more information is needed regarding the Delphi process used. For example, how many individuals participated, what were their general qualifications, how many iterations of the Delphi approach were used before the behaviors of interest were solidified, etc.

Authors response: We thank the reviewer for the positive comments on the Delphi Study. As per the reviewer suggestion, we have added the following information on the number of Delphi rounds, number of participants and their general qualifications: “All the experts involved in the two-round Delphi (round 1 n= 54; round 2 n= 23) had an academic role in universities across Australia and New Zealand and were involved in small-group teaching that included design and/or facilitation of collaborative learning..” (Revised line 131-134, highlighted in yellow)

Adam Wilson (Reviewer 2) comments and suggestions: Given your use of scaled items with anchors being, 'never, rarely, often, and always' I question how many times a student was
evaluated with the CLeD-Ex. These terms suggest that students were longitudinally evaluated, but I did not see evidence of longitudinal evaluation in the study design. Please clarify.

Authors Response: We thank the reviewer for this comment. We would like to clarify that students completed the CLeD-EX once during the course. The items on the CLeD-EX explored the frequency of the observed behaviours in students. The timing of the CLeD-EX instrument completed by the student varied for students in the intervention and control groups. As per the reviewer suggestion, we have updated Figure 1 (revised line 190-196, highlighted in yellow). We have included the time durations and have added notes to clarify that students completed the CLeD-EX once during the course.

Adam Wilson (Reviewer 2) comments and suggestions: Regarding tutor evaluation and feedback, more contextual information is needed about the tutors (i.e., what are the tutors qualifications, are they faculty, are they senior medical students? To what extent were there given frame of reference training for how to use the instrument?).

Authors Response: We thank the reviewer for this comment. We have clarified that the tutors are trained academics in health professions education (Revised line 145-146, highlighted in yellow). In addition, we have also clarified the nature of the training that was provided to the tutors by adding the following statement: “The tutors were introduced to the CLeD-EX through a brief presentation in the course orientation session. The purpose of this presentation was to introduce the tutors to the CLeD-EX and to address any queries about the instrument or the field-test design.” (Revised line 167-169, Highlighted in Yellow)

Adam Wilson (Reviewer 2) comments and suggestions: Figure 1 would benefit from adding time durations to the schematic. For example, how much time elapsed between the 'pretest week 1', the 'CleD-Ex', and the 'posttest week 6'?

Authors Response: We thank the reviewer for this comment. As per the reviewer suggestion we have updated Figure 1 (revised line 190-196, highlighted in yellow). We have included the time durations and have added notes to clarify that students completed the CLeD-EX once during the course.

Adam Wilson (Reviewer 2) comments and suggestions: Under the evaluation section, avoid the use of bullet points and write out the analyses in paragraph format.

Authors Response: We thank the reviewer for this comment. As per the reviewer suggestion this has been amended in text (Revised line 201-218, highlighted in yellow)
Adam Wilson (Reviewer 2) comments and suggestions: I recommend re-running interrater agreement using Cohen's kappa instead of a Pearson correlation. Also consider running a Kendall tau b to help validate the internal consistency of scores generated from the items. Person's correlation is classically used with a continuous scales, your rating scales are ordinal, not continuous.

Thank you for including effect sizes.

Authors Response: We thank the reviewer for their suggestion. We have re-calculated the inter-rater agreement via Cohen’s kappa and run the Kendall Tau-b to validate the internal consistency of items. Please find the updated results in the following section in the manuscript (revised line 265-270, highlighted in yellow)

Adam Wilson (Reviewer 2) comments and suggestions: Results

Regarding feasibility, please indicate how many sessions the tutors completed before providing feedback. Also indicate the number of students tutors observed at a single time and how times students were observed. (This information can be added to the methods and does not need to appear in the results section). Was there any overlap in who the tutors evaluated? It is a bit atypical for self-evaluations to be compared against external evaluators to run an inter-rater agreement analysis. Typically, interrater agreement means that both raters were expert (external) raters, not a self-rater. How strong was interrater agreement between tutors for the same group of students?

Authors Response: we would like to thank the reviewer for these comments. We would like to clarify that the tutors evaluated the student once during the course. Tutors rated the student based on the observed frequency of the behaviour the student exhibited during the course. Each scenario group has one tutor and therefore, a single tutor evaluated the student. As per the reviewer comment, we have added a sentence to clarify the tutor evaluation: “Tutors, while facilitating small group sessions, observe the frequency of the behaviours that students exhibit during the learning activities.” (revised line 145-147), highlighted in yellow). As per the reviewer request to indicate a number of students that tutors evaluated using the CLeD-EX, we have added this information in the manuscript: “Each scenario group included approximately 2-4 students who had volunteered to participate in the CLeD-EX field-test.” (Revised line 243-244, highlighted in yellow).

In addition to this we have also updated figure 1 (revised line 190- 196) and have clarified that CLeD-EX was completed once during the course.

In the analysis we compared the self-evaluation with the tutor evaluation to explore the accuracy of student self-evaluation. We would like to clarify that this study design does not permit the measure the inter-rater reliability between tutors for the same group of students; each tutor was involved in providing evaluation and feedback to the students in their group only.
It was not clear to me how the analysis of student reflections helps to meet the aims of the study. It does not appear to contribute to effectiveness, feasibility/practicality, or impact (which is a nebulous term and should be avoided). As such, I recommend discarding this section. If it is kept, a table should be generated with major themes, minor themes, and representative comments to better streamline this section.

Authors response: We thank the reviewer for their comment on including the student reflections. The students using the CLeD-EX instrument engaged with reflection and the exercise of developing a plan of action for future improvement. The frequency and analysis of the reflective accounts were indicators of students’ engagement with the activity and their commitment to develop their skills in collaborative learning. This is indicative of the effectiveness of CLeD-EX as an educational instrument to promote effective collaborative learning in students. We considered including a table of themes to present the student reflections, however the themes which were discussed in the reflections were related to the behaviours on the CLeD-EX instrument.

The pre-test and posttest evaluation section of the results should be moved up in the results section. Throughout the manuscript, please keep the presentation of findings in the same order as the aims. To me this section speaks to effectiveness of the process as opposed to its 'impact'.

Authors Response: We thank the reviewer for this suggestion. As per the reviewer suggestion we have moved the pre and post-test evaluation results in the results section under the subheading - “Effectiveness and educational impact of CLeD-EX” (Revised line 275 onwards with track changes). In line with these changes the results on the analysis of written feedback on CLeD-EX and student reflection has also been moved under this subheading (Revised line 291 onwards).

In table 4, also report the effect sizes.

Authors Response: In the revised manuscript Table 4 has been changed to Table 8. We note that the Effect sizes have been reported in the manuscript (Revised line 285-287).

Regarding tables related to the McMaster questionnaire that was used, please clarify who completed this questionnaire. Are these scores self-reports from students or did the tutors also fill these out?

Authors Response: As per the reviewer suggestion, this information has been added in the manuscript. “All students participating in the CLeD-EX field-test completed the pre and post-test questionnaire.” (Revised line:222-223, highlighted in yellow)

In the tables, use an * to indicate statistically significant findings.

Authors Response: As per the reviewer suggestion, this has been revised in the manuscript.
Adam Wilson (Reviewer 2) comments and suggestions: Discussion

In the 3rd paragraph, last sentence, you write, "These findings are consistent with the literature which reports that students have a tendency to underestimate their performance during self-evaluation." I can't help but notice you have only one citation for this statement. This is likely because more of the literature has found just the opposite to be the case (i.e., students overestimate their performance). I encourage you to do a deeper dive into the literature to better support this statement with more references or to further explain why your findings don't align with the majority of the literature.

Authors Response: We thank the reviewer for this suggestion. As per the reviewer comments, we have revised and included literature to better support students underestimation and overestimation of self-evaluation of performance: “There is a tendency for weaker students to overestimate their work and for better students to underestimate their performance (43-45). Therefore, a students’ self-evaluation may vary in accuracy. Self-assessment may thus be seen as being of limited value, if our aim is accuracy. However, in the context of developing essential professional skills of self-evaluation, self-assessment could play a significant role.” (Revised line 400-404, Highlighted in yellow).

You note, "only the self-direction scale improved in the control group." You then write, "This finding provides evidence for efficacy of CLeD-EX from the students' perspective." I disagree with this interpretation. To me, this finding suggests that self-directed learning was not a byproduct of the exercise under investigation and was largely unrelated to the construct of interest.

Authors Response: We thank the reviewers for this comment. We have revised the statement in the discussion: “The significant improvement of students on critical thinking and group process subscales in the intervention group provides evidence of the efficacy of the CLeD-EX.” (Revised line 408-410, highlighted in yellow).

Please clearly outline the practical implications of this research in a separate section of the discussion.

Authors Response: We thank the reviewer for the suggestion to include a section on practical implications. Therefore, as per the reviewer request, this section has been added in the discussion: (Revised line 437-448, highlighted in yellow).

Please include a future directions section.

Authors Response: We thank the reviewer for this suggestion. We note that we have included this section with the conclusion section. As per the reviewer suggestion, we have made the future directions section as a separated subheading. (Revised line 455-458)
In the conclusion section, please summary all aims of the study. You mention the CLeD-EX is effective and feasible, but fail to mention anything about its impact, which was also an aim of your study (again I recommend changing the word impact to something which can be measured).

Authors Response: We thank the reviewer for this comment on the conclusion section. Therefore, as per the reviewer suggestion, we have added the statement on the positive impact of CLeD-EX on students’ critical thinking and group processing and the design of CLeD-EX which provided students the opportunity to complete the feedback loop. “Preliminary evaluation of the CLeD-EX indicates its positive impact on students’ critical thinking and collaborative learning behaviours.” (Revised line 451-453, highlighted in yellow)

We look forward to hearing from you regarding our submission. We would be glad to respond to any further questions and comments that you may have.

Kind Regards,

Maha Iqbal, Gary M Velan, Anthony J O’Sullivan and Chinthaka Balasooriya