Author's response to reviews

Title: Example-based learning: Comparing the effects of additionally providing three different integrative learning activities on physiotherapy intervention knowledge

Authors:

Joseph-Omer Dyer (joseph.omer.dyer@umontreal.ca)
Anne Hudon (anne.hudon@umontreal.ca)
Katherine Montpetit-Tourangeau (katherine.montpetit-tourangeau@umontreal.ca)
Bernard Charlin (bernard.charlin@umontreal.ca)
Silvia Mamede (mamede@fsw.eur.nl)
Tamara van Gog (vangog@fsw.eur.nl)

Version: 2 Date: 23 December 2014

Author's response to reviews:

Response to Reviewers

We would like to thank the four reviewers for their thoughtful comments and suggestions, which we feel helped us improve our manuscript. Below, we detail the changes we made in the manuscript in response to each remark.

Reviewer 1
Reviewer's no. 1 report
Title: Learning with case-based worked examples: Comparing the effects of additionally providing self-explanations and making or studying concept maps on physiotherapy intervention knowledge
Version: 1 Date: 30 September 2014
Reviewer: Ralf Schmidmaier
Reviewer's report:
Reviewer no. 1; comment 1
The authors present a well written report on a laboratory experiment with 91 physiotherapy students in an electronic learning environment. The title summarizes the study correctly but is in my opinion only an aggregation of the key words and is not really understandable for the reader. This is good for literature search, but makes it difficult to structure the reading of the article. I prefer to describe the results in the title, not the method, but opinions may differ concerning this point.

Authors' reply to reviewer no. 1; comment 1
Concerns about the title were also raised by reviewer 2 who proposed the following title: “Learning with case-based worked examples: comparing the effects of three integrative activities on physiotherapy intervention knowledge.” We have adopted this suggestion and hope that you agree with this new title.

Reviewer no. 1; comment 2

The abstract is well written and presents all necessary information. The introduction gives a good overview about the previous work of the authors and provides the definitions of the used terms in the manuscript. It tries to explain the operationalization of the study, but neglects to some extent the complexity of the problems (e.g. meaningful learning or problem solving) and avoids mentioning alternatives. The whole story about clinical reasoning is much simplified. The idea of the authors was to combine different instructional methods – one after the other – to find the optimal sequence to foster clinical reasoning maximally. Why scaffolding and reflection and why not other strategies? I assume that these methods fit well to the used instructional method of case-based worked examples.

Authors’ reply to reviewer no. 1; comment 2

We added a part to the Introduction section in which the complexity of clinical reasoning is better explained, and how the acquisition of clinical reasoning can be supported by educational methods that improve knowledge, data gathering and data processing. It is subsequently explained why examples are well suited to foster clinical reasoning, especially in combination with completion problems. (See "Example-based learning" in Introduction lines 61 to 79).

Reviewer no. 1; comment 3

Case based worked examples are an established method to foster clinical reasoning, but there remain many open questions regarding the optimal use – no doubts about this. The authors state that it is unclear how to optimize the acquisition of problem solving skills. What do they mean? Problem solving skills in general? Or rather procedural knowledge in the special field of electrophysical intervention in physiotherapy? Do the authors want to optimize a special course in physiotherapy or do they want to gain knowledge about more general principles of problem solving? And especially in the field of electrophysical physiotherapy we do not know whether the chosen methods are suitable only for novices – like in the study – or also for advanced learners or physiotherapists who lack specific knowledge in electrophysical therapy. I think it is important to discuss these points and to stress the goal of the presented study.

Authors’ reply to reviewer no. 1; comment 3

We have now added a section to introduce readers to the main differences among three types of knowledge: factual, conceptual and procedural knowledge. Moreover, we now detail the components of procedural knowledge and how it relates to problem solving (see " Meaningful learning to foster clinical reasoning" in the Introduction, lines 183 to 205). We have also added details to specify the objectives and operationalization of the study (See "Clinical reasoning in the
The authors state that worked/completion example pairs “may” be an effective strategy for improving physiotherapy knowledge in novice learners and they cite the available evidence. However, is this really proven? However, we have to assume that it is because this question is not addressed in the presented study and this variable was kept stable.

Authors’ reply to reviewer no. 1; comment 4
We have modified this sentence following your comment. To the best of our knowledge, our study is the first to use worked/completion example pairs in health profession education, but several studies have shown the benefits of a fading strategy in which students first study a fully worked-out example and progress via completion examples with increasingly more blanks to independent problem solving. When considering the literature on worked examples that we reviewed in our Introduction, one can state that: “In short, example-based learning with worked/completion example pairs involving backward fading is an efficient learning strategy for fostering meaningful learning of problem-solving skills in novice learners.” (See Introduction, lines 179 to 181).

Reviewer no. 1; comment 5
Based on this the authors want combine this combination with other strategies like self explanation or concept-mapping to promote more meaningful learning …. of procedural knowledge? … of problem solving skills? … of factual knowledge?

Authors’ reply to reviewer no. 1; comment 5
In response to your comment, we now specify in the Introduction that meaningful learning was assessed by problem-solving performance (see Introduction lines 357 to 361).

Reviewer no. 1; comment 6
Meaningful learning in this context means cognitive skills that allow solving new clinical problems – a kind of transfer performance. The authors assume that especially self-explanation would be a potent method to foster (near) transfer. In contrast they assume that concept mapping focuses on conceptual knowledge with less strong effects on transfer performance.

There is a comprehensive review about concept mapping which leads the reader a little apart from the main focus of the study, as concept mapping is a well known method in the literature. The dependent and independent variables are pointed out in the last section of the introduction, but the reason for choosing conceptual knowledge, near transfer and cognitive load as outcome parameters for better clinical reasoning in a special field is not easily to understand.

Authors’ reply to reviewer no. 1; comment 6
We now specify in the Introduction the reasons for choosing conceptual knowledge, problem-solving skills and cognitive load (see Introduction, lines 347-361) as outcome parameters.

Reviewer no. 1; comment 7
The authors clearly show their broad experience in the field and their founded knowledge about the scientific problems and the necessary methods in the introduction. However, I suggest an additional section between meaningful learning and objectives to summarize the rationale, to describe the lack of knowledge, to reason the choice of methods used for operationalization and assessment and also to explain what the limits are and what cannot be addressed by the study in this complex field in clinical reasoning and problem solving.

Authors’ reply to reviewer no. 1; comment 7
We added a section to the Introduction to summarize the rationale and choice of the outcomes (conceptual knowledge, problem-solving performance and cognitive load) (See Introduction, lines 325 to 361). We added in the Discussion a section to explain the limitations and what could not be addressed in this study in the complex field of clinical reasoning (see Discussion, lines 726-742).

Reviewer no. 1; comment 8
In the methods section the authors very nicely describe this complex laboratory study. They try to keep it readable on the one hand, but comprehensive on the other hand. I suggest putting the whole material in a supplemental section, not only one example for each. It would be much easier for the reader (who is interested in the details) to understand the complex interactions between the interventions and the assessments. Conceptual knowledge was assessed by six multiple choice questions. Only 2 questions are presented. These two questions do not fulfill the criteria for good quality MCQs as described in the literature (long statements, cues, several statements within one answer, combinations of answers …). Additionally, I am wondering whether 6 (!) questions of uncertain quality provide reliable results. What are the test statistics for this assessment, please present! Furthermore, I am not sure whether it really assesses conceptual knowledge, many of the given examples are simple textbook facts. Also here, please present the whole test. Problem solving skills were assessed by short case scenarios called transfer test. The authors should use a straight nomenclature, what they really mean, they mix transfer test and problem solving throughout the whole manuscript including the table. They should provide the test statistics and the correct solution of the scenario and the rules how the 20 points were distributed.

Authors’ reply to reviewer no. 1; comment 8
Conceptual knowledge and problem-solving tests (pre-test and post-test), correction grids are provided in Appendix of this rebuttal letter. The reliability of each test is now provided in the Methods section (see "Test tasks" lines 486 and
A main difference between the conceptual knowledge test and transfer test is that one is MCQ and the other is open ended question. Retrieval of knowledge is therefore completely different. As far as I can interpret the content and the knowledge dimension of the conceptual test and the transfer test it is almost the same and both could have been assessed by the same type of question (either open ended or MCQ). Only the difference in retrieval of knowledge (active vs. passive) may explain the differences of the study. I suggest being very careful in the interpretation of the results and generalization of the obtained results in terms of conceptual knowledge, problem solving skill and transfer performance.

Authors’ reply to reviewer no. 1; comment 9

We agree that the difference in the type of retrieval between the conceptual knowledge and problem-solving tests may have influenced the results of this study. We now mention this in the interpretation of the results in the Discussion (see lines 726-735).

Reviewer no. 1; comment 10

It would have been interesting to change the sequence of meaningful learning methods to find out, why self explanation was so useful. It can be assumed that self explanation is most potent at the beginning of the session with subsequent studying units.

Results: the authors present theirs results in a very well manner, scientifically correct. I am not surprised by the fact that the concept map study group outperformed in the rather factual knowledge based test as they were confronted with the most conceptual facts during the study. They spent more time on facts studying than the other groups that were prompted to organize knowledge. Of note, there was no increase of knowledge between pre-test and post-test in the latter two. I am also not surprised that the self explanation group outperformed in “problem solving”, because the open ended prompt in the self explanation task is quite similar to the open ended questions in the problem solving tests (“to treat a patient with any pain related problem” vs. “most appropriate intervention”).

Authors’ reply to reviewer no. 1; comment 10

All participants performed three completion examples in which they had to answer open-ended questions during the learning phase that were the same as those asked in the problem-solving pre-test and post-test. These questions were formulated in the learning phase and tests as follows: 1) “What is the most appropriate electrotherapy intervention for this case?”; 2) “What are the optimal adjustment parameters for the electrophysical agent selected?” and 3) “What key characteristics of the case justify the intervention selected?” At the end of the learning phase, all participants of this study were familiar with answering these questions, which were presented to them in the post-test as well.
Reviewer no. 1; comment 11
Unfortunately the study has not included a control group (no fostering of meaningful learning). Is studying concept maps superior to control? Furthermore, the study focuses on very short term effects (within the same session without distractor). Meaningful learning would mean that near transfer (or far transfer) cases could be solved in a long term manner. Therefore, this aspect of the design of the study seems to be crucial for the interpretation of the results. Furthermore, facing the fact of content specificity of knowledge and clinical reasoning, the question whether similar results are obtainable in completely different content domains is not addressed.

Authors' reply to reviewer no. 1; comment 11
We now address these relevant comments in the Discussion (see lines 744-766)

Reviewer no. 1; comment 12
Again I have a problem in reading the first section of the discussion as transfer and problem solving are mixed up. Please use a uniform nomenclature to make the manuscript better understandable. To some extent the study design includes repeated testing (e.g. self explanation and completion cases) and repeated studying (concept map study group). These issues should be brought into consideration.

Authors' reply to reviewer no. 1; comment 12
We now use the terms "problem-solving" throughout the manuscript to be consistent in terminology and avoid confusion. We added some additional text to consider the fact that this study involves repeated exposure to the learning activities and how this could have influenced the results. (see lines 744-766)

In summary This is a quite large, prospective, well controlled, randomized trial regarding an important topic in medical education The manuscript is well written and the data are presented in a usual scientific manner The study design is well described with clear dependent and independent variables and interventions Interventions are clearly described and well done

Reviewer no. 1; comment 13
Operationalization and nomenclature are not stringent. This should be revised. Knowledge gap and rationale should be very clearly stated.

Authors' reply to reviewer no. 1; comment 13
Operationalization, knowledge gap and rationale: please see our response to comments 3, 4, 5, 6 and 7. Nomenclature: please see our response to comment 12

Reviewer no. 1; comment 14
The assessments are rather weak, so interpretation should be careful. Limitations of the assessment should be honestly and clearly described in the introduction, in the methods section and especially in the discussion. The perfect assessment of problem solving skills is difficult, so alternatives should be discussed and it should be clearly stated that only a (small) piece was assessed in this study. Interpretation of the results should be more descriptive, especially in the results section, in the abstract and title. Generalizations should be clearly named as hypotheses. The limitations of the study should be discussed more intensively.

Authors’ reply to reviewer no. 1; comment 14
Please see our response to comment 8

Reviewer no. 1; comment 15
As the topic is important, the study is large and well done I recommend publication. However, the mentioned aspects with nomenclature, operationalization and data interpretation need to be addressed before publication (major revision).

Level of interest: An article of importance in its field
Quality of written English: Acceptable
Statistical review: Yes, and I have assessed the statistics in my report.
Declaration of competing interests:
I declare that I have no competing interests

Reviewer 2
Reviewer’s report
Title: Learning with case-based worked examples: Comparing the effects of additionally providing self-explanations and making or studying concept maps on physiotherapy intervention knowledge
Version: 1 Date: 1 October 2014
Reviewer: Thomas Jaarsma
Reviewer’s report:

Minor Essential Revisions

Reviewer no. 2; comment 1
1. Overlap between intervention and post-test
This study shows interesting effects of an integrative learning activity on both conceptual knowledge and problem solving skills. However, there seems to be quite an overlap in the tasks in the condition and those in the post-test. The questions asked in the self-explanation condition, for example, resemble those in
the post-test on problem solving skills to quite some extent. The same goes for the concept map study condition and the knowledge questions in the post-test. How surprising, then, are the results obtained? Put differently, are there grounds to expect that studying a model concept map could by some means increase problem solving skills at all?

Authors' reply to reviewer no. 2; comment 1

In the Introduction ("Hypotheses" section) we added some evidence that suggests that concept mapping can foster problem-solving skills (see Introduction, lines 305-308). We also added some arguments to explain how concept map study might foster more learning than concept mapping and self-explanation by reducing the cognitive load during learning (see Introduction, "Hypotheses," lines 311-318).

To more specifically reply to your concern about intervention and post-test:
Questions in the problem-solving tasks of the pre-test and post-test were purposely the same as those of the completion examples but different from the self-explanation prompts and from the questions of the concept mapping and concept map study conditions. Indeed all participants (regardless of the learning condition) were asked to answer the following questions in both the completion examples and problem-solving tasks in the pre-test and post-test: 1) "What is the most appropriate electrotherapy intervention for this case?"; 2) "What are the optimal adjustment parameters for the electrophysical agent selected?" and 3) "What key characteristics of the case justify the intervention selected?". Therefore all participants were familiar with answering these questions when they performed the post-test. However, it should be noted that this was not sufficient, as the tasks in the post-test had different surface features. So students could not just repeat the answers they had given in the learning phase; they had to understand the problem at hand, and for that, concept mapping or map studying could be helpful.

Reviewer no. 2; comment 2
2. Structure of the introduction

The background section in the abstract illustrates what – in my opinion – could be improved in the article’s introduction: every sentence in that part of the abstract introduces a new instructional/learning method. These concepts are all discussed in the introduction, too. Are all these steps necessary to come to the objective of this study? I think it poses a large stress on the reader to follow this line of reasoning throughout the introduction and to get where the article is going: the effect of integrative learning activities.

Authors’ reply to reviewer no. 2; comment 2
To ease the reading, Abstract and Introduction sections have been rewritten. The background section of the Abstract has been modified (see Abstract lines 2-13). We added a new section in the Introduction section (first paragraph) to give readers a concise overview of the background and of the aim of the study before
getting into the specific descriptions of the learning activities used in the study (see Introduction, lines 41-59). Other parts of the Introduction have been rewritten according to other reviewers' comments.

Reviewer no. 2; comment 3
3. Introduction of the domain

An additional comment on the introduction is the lack of introduction and justification of the domain (physiotherapy). For example, the case study of electrophysical agents is now introduced in the ‘Participants and Design’ paragraph; this might better fit in a short paragraph on the domain in the introduction?

Authors' reply to reviewer no.2; comment 3
We have now included a section about the domain in the introduction: "Clinical reasoning in the physiotherapy domain.".(see Introduction, lines 325-361).

Reviewer no. 2; comment 4
4. Justification of problem solving skills and conceptual knowledge

As it appears to me, clinical reasoning (the main topic of the article) is operationalized by the constructs problem solving skills and conceptual knowledge. However, a clear justification why clinical reasoning boils down to these two constructs seems to be lacking. Their first appearance is in line 119, where it is stated that self-explaining can benefit conceptual knowledge. After that, there is no such justification given.

Authors' reply to reviewer no.2; comment 4
The Introduction has been rewritten to explain why clinical reasoning has been assessed by conceptual knowledge and problem-solving skills. In a new section "Meaningful learning of clinical reasoning" we describe how we operationalized meaningful learning of clinical reasoning through acquisition of problem-solving skills. In this section, we also explain the relevance of conceptual and procedural knowledge in clinical decision making (see Introduction, lines 190-205). In another new section "Clinical reasoning in the physiotherapy domain" we explain the link between procedural knowledge and problem-solving skills and how we operationalize problem-solving skills in this study (see Introduction, lines 325-361).

Reviewer no. 2; comment 5
5. Inconsistency of terms used

Throughout the article, different terms are used for the same concepts. I believe that the authors mean the same thing by ‘transfer’ and ‘problem solving skills’. In the research question in lines 174-175, transfer is used, while in line 177 ‘problem solving skills’ is used. Am I correct or confused? (this paragraph might need some further attention: two short questions in a row (lines 174-176) and two similar statements in a row (lines 176-178)). The same goes for ‘model concept map condition’ in line 268; throughout the rest of the article, ‘concept map study
Authors’ reply to reviewer no.2; comment 5

The manuscript has been revised thoroughly to avoid the use of different terms for the same concept. The terms “transfer” and "model concept map condition" were replaced by "problem-solving" and "model concept map study condition" respectively.

Reviewer no. 2; comment 6
6. Values in Table 1
To me, it was not clear what the values in Table 1 represent, or what the maximum scores were.

Authors’ reply to reviewer no.2; comment 6
Following comments made by reviewer no. 3, data in Table 1 is now presented in two Tables (new Table 1 and Table 2) and maximum scores are presented.

Discretionary Revisions
Reviewer no. 2; comment 7
7. Concept mapping condition
When discussing the hypotheses (lines 173-187), concept mapping is expected to be less effective than self-explanation. A thought that popped up in my mind: why include it? To test this hypothesis (/expectation), obviously, but is there any positive reason for including concept mapping? Otherwise, any strategy could be included, just to prove it was less effective… What do the authors expect from that activity?

Authors’ reply to reviewer no.2; comment 7
In the new "Hypotheses" section in the Introduction, we added some explanations about positive reasons for including concept mapping in the study (lines 296-323). Some evidence in health profession education show that concept mapping can lead to significant improvements in understanding and problem solving when compared to traditional teaching methods [1-3]

Reviewer no. 2; comment 8
8. Title
I think the title includes too much detail, making it a bit long and incomprehensible. My suggestion would be to substitute ‘additionally providing self-explanations and making or studying concept maps’ by ‘three different integrative learning activities’, e.g.

Authors’ reply to reviewer no.2; comment 8
We agree with this comment. The title of the manuscript has been changed accordingly.

Reviewer no. 2; comment 9
9. Participants and design section.

As I said earlier, I think this section deserves some attention. It now includes information on the task, while information on which conditions are present in the study (which I would expect in a description of the design) is lacking. Also, I would move the comments on the ethics (lines 204-206) to the back as they now interrupt a description of the participants’ prior knowledge/training.

Authors’ reply to reviewer no.2; comment 9

We have now included information on which conditions are present in the study (see Participants and design section, see lines 384-402). Comments on the ethics have been moved to lines 374-375.

Reviewer no. 2; comment 10

10. Data on switches between worked and completion examples?
Just a loose thought: In lines 319-320, it is mentioned that participants can switch as they wish between the worked and completion example. Is there data on the number of these switches? Perhaps that more motivated participants would switch more back and forth to optimize their completion task and hence learn more/better...? Put differently: could it be a covariate?

Authors’ reply to reviewer no.2; comment 10

Unfortunately, the digital learning environment was not designed to allow us to evaluate the number of these switches. We now discuss this as a limitation and how adding another assessment such as eye-tracking might be relevant in testing whether more switches are associated with better learning outcomes—or worse learning outcomes, for that matter, since multiple switches back and forth might also indicate that learners did not learn sufficiently from the example to solve the steps in the completion problem (see Discussion lines 707-724).

Reviewer no. 2; comment 11

11. Discussion on far transfer

Self-explanations are mentioned as likely contributors to far transfer. What about concept mapping? This activity has a high cognitive load, but I could imagine that this leads to a better retention on the long term (as compared to simply studying a complete concept map, e.g.).

Authors’ reply to reviewer no.2; comment 11

We now discuss the potential of concept mapping in fostering far transfer (see Discussion, lines: 768-774).

In the present study, near transfer refers to learners’ capacity to solve problems that have different superficial characteristics but share similar deep structure than the problems studied. In contrast, far transfer refers to learners’ capacity to solve problems that have different superficial characteristics and deep structure than the problems studied. Some evidence suggests that studying model concept maps as advanced organizers can foster far transfer [4-6]. When considering
learners’ constructing concept maps by themselves (i.e., concept mapping), several studies have shown that concept mapping can promote near transfer in health profession education [1-3]. However, there is no evidence, to the best of our knowledge, that concept mapping is better than other learning activities to foster far transfer.

When considering delayed post-tests, it is still unclear which activity between concept mapping and concept map study is better for delayed recall. Both concept map study and concept mapping can improve delayed recall compared to other learning activities.

Level of interest: An article of importance in its field
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.

Reviewer 3

Reviewer's report
Title: Learning with case-based worked examples: Comparing the effects of additionally providing self-explanations and making or studying concept maps on physiotherapy intervention knowledge
Version:1 Date: 2 October 2014
Reviewer: Martine Chamberland
Reviewer's report:
- Major Compulsory Revisions
None
- Minor Essential Revisions
Reviewer no. 3; comment 1
Throughout the text, the authors seem to use problem-solving skills and transfer performance as synonyms. This might be confusing for the reader since the authors did not make that equivalence explicit in the text and I am not sure that in fact the two terms are interchangeable. Transfer being usually defined as the ability to extend what has been learned in one context to new contexts.

Authors’ reply to reviewer no.3; comment 1
The manuscript has been revised thoroughly to avoid the use of different terms for the same concept. The terms "transfer" and "model concept map condition" were replaced by "problem solving" and "concept map study condition" respectively.
Methods:
In the paragraph under «Experimental procedure», it is mentioned that participants completed a mental effort rating scale after the pre-test. I don’t clearly understand the rationale of doing this measure at that point (unless it was an opportunity for the participant to familiarize with the scale); moreover, the results are not reported later.

Authors’ reply to reviewer no.3; comment 2
This was a mistake. Mental effort was not evaluated during the pre-test, but during the learning phase and post-test. This sentence has been deleted from the Experimental procedure section.

Reviewer no. 3; comment 3
Did the participants have any specific preparation or instructions for self-explanation? Please clarify this point.

Authors’ reply to reviewer no.3; comment 3
Participants did not receive any training for self-explanation, but they did receive brief instructions about how engaging in self-explaining can foster learning (the concept map conditions were given similar information re. concept mapping/concept map study). This is because we wanted to use a prompting intervention (not a training intervention) to foster self-explanation in a digital learning environment similar to that used by Atkinson et al., 2003.[7, 8]. We added some theoretical explanations in the self-explanation section of the introduction (see Introduction, lines 222-231)

Reviewer no. 3; comment 4
Data Analysis:
To keep the line of thinking obvious for the reader, I would suggest to present first the analyses used to answer the first question (learning outcomes of the conditions) and second, the analyses used to answer the second question (mental effort).

Authors’ reply to reviewer no.3; comment 4
We agree this would improve the readability. The Data Analysis section has been modified accordingly (see Data Analysis, lines 550-569)

Reviewer no. 3; comment 5
Results:
Table 1 is too busy and shows different types of outcomes. I would suggest to present the results in two different tables:

Table 1: Conceptual knowledge and problem-solving; and Table 2: Mental effort
In the results of the pre-test, since there is no significant difference in the
ANOVA, I think that Tukey post-hoc analyses could be deleted.

Authors’ reply to reviewer no.3; comment 5

Data in Table 1 is now presented in two tables (new Table 1 and Table 2). Maximum scores are presented.

Reviewer no. 3; comment 6

Again, to keep the line of thinking obvious for the reader, the authors should present first, the results of the first question and second, the results of the second question.

Authors’ reply to reviewer no.3; comment 6

The results section has been reorganized accordingly (see Results, lines 573-611).

- Discretionary Revisions

Reviewer no. 3; comment 7

Background :
I am not sure that the section entitled « Meaningful learning » is absolutely necessary. Also, since it introduces the following sections on self-explanation and concept maps, it gives the impression that worked examples described in the previous section do not lead to meaningful learning, which is not, I think, the case.

Authors’ reply to reviewer no.3; comment 7

The Introduction has been reorganized according to other reviewers’ comments. This should resolve this issue.

Reviewer no. 3; comment 8

Discussion :
If participants did not have any specific preparation for self-explanation, this point could be discussed as an additional advantage of this strategy.

Authors’ reply to reviewer no.3; comment 8

In the Discussion section, we now mention the additional advantage of self-explanation when considering the fact that learners had no specific training for this technique (see Discussion, lines 644-648)

Reviewer no. 3; comment 9

Appendices : I think there are too many appendices. I would keep only appendix A, C and D

Authors’ reply to reviewer no.3; comment 9

We have decided to keep all appendices according to reviewer 1’s comment (no.8.)
The authors describe an empirical study that investigates the effects of three different scaffolds that are added to a digital, case-based learning environment designed to foster physiotherapy students' acquisition of conceptual knowledge and problem-solving skills (transfer): After having studied a worked example/completion example pair (overall, three such pairs were given), students in the first experimental condition were prompted to self-explain the solution steps of the examples (self-explanation condition). Students in the second experimental condition received an incomplete concept map and were instructed to complete it after each example pair (concept mapping condition). Students in the third condition received an already completed concept map that represented the main concepts and relation between concepts that were covered in the examples (model concept map study condition). Dependent variables were the mental effort students reported during the process and the pre- and posttest (i.e., when being exposed to the worked examples/completion problems), the conceptual knowledge students and the problem solving skills (i.e., transfer performance) they gained from pre to post test. Results indicated that conceptual knowledge acquisition was best fostered in the study concept maps condition (whereas the other two conditions were less effective, but did not differ from each other), whereas problem-solving skills were best acquired through self-explaining, whereas students from the two concept mapping conditions showed significantly lower performance (but did not differ from each other). Data from the mental effort instrument indicated that mental effort was lowest in the concept map study condition during the learning process, whereas during post test it was lowest in the self-explanation condition and generally negatively correlated with post test performance. Overall, the authors argue that being exposed to self-explanation prompts is likely to promote the construction and automatization of schemata (i.e. promote understanding) that could then easily
be used during post test. Concept mapping, in contrast, might have provided learners with an opportunity to incorporate in-depth concepts into their schemas (i.e., promote information chunking), without necessarily understanding the relations between the different concepts that were presented/created in the concept maps. Overall, the manuscript meets high scientific standards. The theoretical part displays a thorough and concise line of argumentation that nicely prepares the research questions and hypotheses. The method of the empirical study is sound and follows the typical experimental paradigm (including the instruments that were selected) that is often used in research on the effectiveness of (different variants of) worked examples. The results of the empirical study are presented in a reasonable way and the discussion of the findings is compelling and well-grounded in theoretical arguments that were previously presented and developed in the theoretical part. Having said this, my following criticisms are not at all meant as reasons for a rejection of the manuscript. Rather, they should be seen as constructive criticism that might strengthen some of the parts of the paper:

Reviewer no. 4; comment 1

1. In the theoretical section, I felt that a bit more justification could be given for why case-based learning is supposed to be a helpful instructional approach in medical/physiotherapy education. Especially from the cognitive load theory perspective that the authors are taking, it seems a bit odd to promote an approach that proponents of cognitive load theory would probably criticize for evoking high levels of extraneous load (due to the fact that students would have to find solutions for the presented cases themselves – at least if case-based learning would be unstructured). Having said this, the authors might want to consider extending the arguments in favor of case-based learning that now only cover the bottom part of p. 4 and perhaps even devote a whole section to case-based learning that provides arguments for the approach. Arguments for case-based learning can be derived from work by Kolodner or Schank.

Authors’ reply to reviewer no.4; comment 1

We have now revised the introduction and have included a section dedicated to "Example-based learning" where we discuss learning from (worked-out) cases only in the context of example-based learning, because that is the most relevant for our study (not case-based learning in general). (see Introduction, lines 61-124). We also now refer to "Example-based learning" instead of "Case-based learning" throughout the manuscript.

Reviewer no. 4; comment 2a

2. Although the presentation of cognitive load theory is very concise, it might be a bit more differentiated. Recently, Ton de Jong and Slava Kalyuga have expressed criticism regarding various aspects of CLT (e.g., regarding the question whether there are really three or perhaps only two types of CL).

Authors’ reply to reviewer no.4; comment 2a
We have now included a short description about the recent debate on whether there are three types of cognitive load (see Introduction, lines 146-160).

Reviewer no. 4; comment 2b

Also, the argument that “worked examples facilitate learning by decreasing cognitive load, and increasing germane load: all available working memory capacity can be dedicated to schema construction” might be complemented with explicitly stating that however it is not guaranteed that students will automatically invest working memory capacity in learning processes (germane load). Actually, making this open question explicit would nicely fit to why you added scaffolds to the worked examples. I.e. you could argue that to increase the likelihood that worked examples work (i.e. that students really use the opportunity of a decrease in extraneous load to engage in activities that “represent” germane load), further scaffolds/reflection prompts are necessary.

Authors’ reply to reviewer no.4; comment 2b

We now explicitly state in the Introduction section that presenting worked examples does not guarantee in itself that learners will effectively use the working memory capacity available for learning processes. Indeed, as the reviewer indicates, this also resonates well with our use of completion problems and additional instructions or activities to ensure that learners are more actively engaged in studying worked examples (see Introduction, lines 164-166).

Reviewer no. 4; comment 3

3. One aspect I really like about the theoretical part is that it presents quite differentiated arguments for why self-explanation prompts would work (and for what; see top of p. 8: “The key aspect of the self-explanation effect is that learners attempt to revise their understanding and make sense of the material, even if they are unsuccessful in articulating a correct explanation.”), and why concept mapping might work (and for what; see p. 9: “Concept map study may (...) foster deep learning by encouraging learners to judge the importance of the concepts and process their hierarchiacal organization and the relationships between them.”). Interestingly, these are not arguments that come from CLT, but rather from other cognitive approaches. Yet, in the empirical study, only a CLT-related measure is taken as a process measure. Thus, the lack of further process data that would for example more directly test the assumptions just mentioned (e.g., think-aloud data, eye-tracking data etc.) might be highlighted as a limitation of the study and a desiderate for future studies.

Authors’ reply to reviewer no.4; comment 3

We now indicate as a limitation of the study that process data are lacking and explain why future research in this area could use log file data (see reviewer 2’s point about switches), think-aloud, or eye-tracking data to achieve a better insight into how learners process the examples and tackle the explanation and concept map (study) activities. (see Discussion, lines 707-724).
Reviewer no. 4; comment 4

4. Another criticism that might be highlighted in the limitation section refers to the rather low external validity of the study. Based on Chi’s (2009) ICAP model, for example, one might hypothesize that having students create concept maps will be more effective (especially regarding transfer) than having them self-explain or study concept maps, because the former is a more constructive learning activity than the latter which – according to the data Chi has collected from other studies – is more conducive for learning than what she calls “active” activities (a category into which both self-explaining and studying concept maps might fall into). However, this will probably only hold in settings where students do not have time pressure and can devote as much energy to the task as they want. In your study, however, the time for the single steps was limited, which may have been a major reason for why concept mapping has led to less positive results on transfer than self-explaining.

Authors’ reply to reviewer no.4; comment 4

We kindly disagree that this would be effective for novices even if they had more time. Similar arguments have been made regarding problem solving compared to worked example study (i.e., making errors would be beneficial for learning) and research evidence accumulated over the past years that this is not at all beneficial for learning when students lack prior knowledge or effective strategies. Once some prior knowledge is available, however, it indeed becomes likely that students start to benefit from such constructive activities. Therefore, it is likely that in extended periods of learning, concept mapping may become more effective as learners gain knowledge. We now address this in the discussion (see Discussion, lines 744-766).

Reviewer no. 4; comment 5

5. In the methods section I was a bit surprised by the fact that all students received training in concept mapping, but that there was no training in how to give self-explanations. Through providing learners with a concept mapping training, students from the two experimental conditions that creating resp. studied concept maps might have been advantaged over the students from the self-explanation condition, since the former were already familiar with the scaffold they would receive later, while the latter were not. The authors should justify this decision.

Authors’ reply to reviewer no.4; comment 5

We have now added some information in the theoretical background to make a clear distinction between self-explanation promoted by training and prompting (see Introduction, lines 222-231). In this study, we use prompts to induce self-explanation rather than training, although students did get a brief introduction about how self-explaining can foster learning. This choice was based on previous studies in which this method was used in a digital learning environment [7, 8].
Reviewer no. 4; comment 6
6. Also in the methods section, please explain whether students were forced to write something on the self-explanation sheets or whether they were free to do so. In both cases, I would be interested in how extensively the self-explanation sheets were used (i.e., how much students wrote).

Authors' reply to reviewer no.4; comment 6
As explained in our previous comment, learners were not "forced" to write something in self-explanation text boxes, but they were strongly encouraged to do so (see Method lines 391-394). In fact, all students in the self-explanation condition wrote something in the self-explanation text-boxes. However, three students provided very poor written self-explanation (no more than a few written words per text box). However, one should take into account that written self-explanations do not necessarily reflect some self-explanations that students might have performed mentally.

Reviewer no. 4; comment 7a
7. In the concept mapping condition, I was surprised that students received incomplete maps instead of blank sheets of paper on which they would be asked to create concept maps from scratch. I understand that this probably has to do with the fact that students in the concept mapping condition would not be able to complete this task within the 15 minutes that were assigned to it (by the way: did students in all conditions use the 15 minutes in a comparable way, i.e. was actual time on task comparable across conditions?).

Authors' reply to reviewer no.4; comment 7a
Yes, actual time on task was comparable across all conditions. The interface was designed so that all students had to use all the time allowed to perform each of the learning activities. For the additional integrative learning activity, interfaces for additional self-explanation, completing concept map and concept map study were presented to learners for 15 minutes. However, our design did not allow us to assess how actively students were engaged in these tasks during the allocated time. In this respect, eye-tracking data or think-aloud data might be useful tools to assess learners’ engagement in example-based learning in future studies (see Discussion, lines 707-724).

Reviewer no. 4; comment 7b
Then, however, the question is whether the label “concept mapping condition” actually transports what happened correctly, or whether perhaps “completing concept maps condition” or “incomplete concept maps condition” or something like this would be more accurate.

Authors’ reply to reviewer no.4; comment 7b
We agree that would be a better term and have now labelled this condition as "concept map completion."
Reviewer no. 4; comment 8
8. One problem that is inherent to the comparison of prompting for self-explanations on the one hand and producing/studying concept maps on the other is that students in the two concept mapping conditions were simply presented with more learning material (the information that was presented on the incomplete/complete concept maps) than students from the self-explanation condition. That way, higher post test performance of the study condition on conceptual knowledge may simply be caused by pure memorization effects. In other words, memorization might be an alternative explanation to the nice arguments you provided in the theoretical part on what cognitive processes may lead to the result that concept mapping should work (see point 3 in my review). Also from this point of view, it might be interesting to look for further process measures to find out what processes actually were responsible for the results. If no such process measures were used, the authors should at least speak to this problem in the limitations section.

Authors’ reply to reviewer no.4; comment 8
We agree that this is a potential limitation of the study. We now discuss it as such (see textual comment in manuscript Discussion, lines 657-661).

Reviewer no. 4; comment 9
9. As most research from the cognitive load paradigm, this study only uses a general measure of mental effort which only gives the opportunity to interpret post hoc on what kind of cognitive load has been promoted in the different experimental condition (a criticism that has also been expressed by de Jong). Although using this undifferentiated instrument still seems to fit to the zeitgeist of cognitive load research, I would suggest to critically discuss what other opportunities might be used to measure cognitive load in a way that would allow to differentiate between the three (or two) different load types.

Authors’ reply to reviewer no.4; comment 9
We agree with this comment that the subjective measure of mental effort used in this study only gives an a posteriori general measure of cognitive load. We now discuss about other self-reported measures that could have been used to differentiate between the different types of load, although evidence that these instruments actually do distinguish among the three types of load is not very strong. We also mention objective measures as a possibility (see Discussion, lines 695-705).

Reviewer no. 4; comment 10
10. Information on (interrater) reliabilities might be moved from the section “data analysis” to the section on “test tasks”.

Authors’ reply to reviewer no.4; comment 10
The information on (interrater) reliabilities has been moved from the section “data
analysis” and included in the section on “test tasks” (see lines 504-507).

Reviewer no. 4; comment 11

11. Minor issue: On p. 17, I would explicitly mention what you used as a covariate in your ANCOVAs (pretest performance).

Authors’ reply to reviewer no.4; comment 10
We now explicitly mention what we used as a covariate for ANCOVAs (see Data Analysis, lines 560-562)

Reviewer no. 4; comment 12

12. Given the results, one might dare to hypothesize that a combination of concept mapping (esp. studying) and self-explanation prompts would be more effective than each of the two scaffolds alone. This might be stated as a hypothesis that might be investigated in future research.

Authors’ reply to reviewer no.4; comment 12
We agree that this would be an interesting hypothesis to explore in future studies. We now mention this in the Discussion (see Discussion, lines 675-682).

Overall, I very much enjoyed reading this paper and - in light of my criticisms above – would argue for minor revisions before the manuscript might be accepted for publication in BMC Medical Education.

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