Reviewer’s report

Title: How the Study of Online Collaborative Learning Can Guide Teachers and Predict Students’ Performance in A Medical Course

Version: 0 Date: 30 Mar 2017

Reviewer: Ingo Kollar

Reviewer's report:

The present manuscript describes how social network analysis (SNA) can be used to uncover the nature of social learning processes within communities of learners as well as the impact of these uncovered processes on individual achievement. The study was run in an online surgery course for Medical students in which they engaged in clinical case discussions. Using a web-based online forum, students had the opportunity to exchange messages with each other and the instructor. Social network analysis was used to (a) visually represent the flow of the discussions, (b) to extract a variety of calculated network parameters and (c) to use these parameters as correlates and predictors for student achievement. The results showed that the visual representations that were created led to the identification of interesting discussion patterns both on an individual level as well as at the course level. Also, it appeared that some, but not all of the extracted parameters seem to be of importance (predictive value) for student achievement. While a considerable amount of variance in student grades could be explained by a conglomerate of all the extracted SNA parameters, it seemed that only information centrality and Eigenvector centrality were significant predictors of student grades. Additional correlational analyses yielded a couple of more relations between network metrics (e.g., in-degree) and centrality measures (e.g., Closeness centrality) with student grades. Overall, the study is innovative and demonstrates the potentials of using a still somewhat novel analytical technique to gain a better understanding of social learning processes that are occurring within communities of learners.

Thus, while I am generally positive about the manuscript, I believe that several aspects need to be worked out to recommend if for publication in Medical Education. More specifically, I have two major and a couple of minor concerns:

Major concerns:

1. My most important criticism refers to the question "What exactly do we know more after having read this manuscript?" While I do see the potentials of SNA to uncover processes that may not as easily be observable with other techniques, I am wondering whether this actually really is "big news". As the authors themselves describe, there is already plenty of research available that has used SNA for similar purposes (granted, not in Medical Education though)
that has demonstrated the potentials of the technique. In large parts of their own empirical study, it seems that the already known potentials of SNA are just "demonstrated once again". Where it gets more interesting, though, is when the authors look at the correlations and predictive value of the different network parameters with and for student achievement. There, it becomes obvious that with SNA, the researcher is becoming capable to rather quickly (i.e. without cumbersome and time-consuming content analysis, for example) identify aspects of the social learning process that actually matter for learning. To me, this leads to recommendation that it would be good to shift a bit the focus of the analysis - away from "demonstrating the already known" and towards the impact that different aspects of the learning process that have been uncovered through SNA have on achievement or perhaps further criterion variables the authors might have available. For example, I was wondering whether it would be possible to check whether the two discussion threads that are described on p. 18 have contributed equally or differently to student achievement. Perhaps it would also be possible to include a somewhat less general criterion variable than student grades (e.g., point scores for particular aspects of knowledge etc.)?

2. Related to this, I found one point in the authors' line of argumentation rather weak: Repeatedly across the manuscript, they argue that data that can be obtained from SNA are likely to be more closely connected to learning than "attributes such as age, gender, and disposition" (p. 7). I wonder whether comparing the effects of parameters that can be obtained via SNA to the effects of learning prerequisites on learning outcomes actually is (a) fair and (b) informative. Of course, what is the closest to learning outcomes will always be the processes that preceded these outcomes, not some very distal and general predispositions of variables such as gender or age. A fairer and more informative comparison would be one that looks at the predictive value of SNA parameters on the one hand and process indicators that may be derived from (admittedly more complex and cumbersome, but on the other hand semantically more rich!) content analytical approaches, because both of them refer to the same entity: the learning process. Having said this, I would recommend to make this comparison a more central one, at least in the theoretical part and in the discussion. It might even be feasible to use the available data to run such a content analysis and compare the results of that to the results of the SNA. Yet, I do not see it as absolutely necessary to run such a cumbersome analysis to make the point I just described. Thus, it should be enough to address it in the theoretical part and the discussion section.

Minor concerns:

1. Large parts of the theoretical part and also of the results section read a bit like a "commercial" for SNA. This impression is partially due to the fact that large parts of the manuscript seem to treat SNA as some kind of end in itself, but is at least partially coming from sentences that attribute "inherent power" to the technique (like it would be that once
you choose SNA, your job as a researcher is done). E.g., on p. 4: "Social network analysis can analyze the social structure of a course..." - I would try to change such wordings into: "By aid of social network analysis, researchers are empowered to analyze the social structure of a course...". There are many instances in the manuscript that are similar. These should be changed. (another example: p. 15: "The first sociogram... demonstrates the ability of SNA to summarize all course interactions").

2. In the background section, at several instances I found statements that were too strong for my taste. E.g., on p.5: "It (TEL) also facilitates networked learning by means of computer-supported collaborative learning - features that have been proven to positively and effectively enhance learning". This sentence gives the impression that once you use TEL, everything will be fine. In contrast, a lot of CSCL research has shown how important it is to design TEL in a meaningful and purposeful way to really support learning. Thus, this statement should be revised. Also, it would be good to give a couple of references to CSCL research on scaffolding and scripting collaboration that illustrates this point (e.g., the work by the group around De Wever in Ghent or Fischer/Weinberger in Munich).

3. P.5: "Academically, it (collaborative learning) simulates skills of critical and higher-order thinking... and can be used as an alternative method for evaluation and feedback" - the first part of the sentence again seems to be too general to me (there is plenty of descriptions in the literature that show that collaborative learning can also become horribly wrong!); the second part I do not understand. How can collaborative learning become an alternative method for evaluation and feedback?

4. P.6, lines 31-41: Here you seem to juxtapose LMSs and SNA by saying that LMSs offer "limited insight into studying interactions", while SNA offers more of such insight. I am wondering whether these two concepts are on the same page. I would rather see it like that: LMSs can give you the raw data that you need to be able to run analyses such SNA. So the former is more like a prerequisite for the latter. Or in other words: LMSs are not an analytical, but rather a pedagogical tool. SNA is an analytical tool.

5. P. 7, line 49: Here you refer to the "importance" of a student for the discussion. It would be necessary to already know here how importance is determined (as you could also determine it in a content-analytical kind of way by rating how qualitatively high- or low-ranking that students' utterances are content-wise).

6. I did not completely understand all of the different SNA parameters that are introduced on p. 8 to 10. E.g.: "The harmonic centrality index is very close to closeness centrality index but better suited for disconnected graphs". This description still does not tell me exactly what "harmonic centrality index" actually is. The same a bit later when you talk about betweenness centrality: "It is similar to information centrality, which is another measure of a node in information flow and is an important measure of learners' networks" - so what IS information
centrality then exactly? I actually found the examples you gave on p. 10 helpful to get a more concrete picture of what at least some of the indicators mean and what they can be used for. I would thus recommend to add examples and research objectives to the descriptions of each of the indicators (instead of naming some of these only after you have introduced the indicators).

7. p. 10: I would like to read more about the context of the study by Sie et al. - otherwise it is hard to understand sentences like this: "Their model showed that both powerful and less powerful innovators would benefit from the recommendation system..." - what exactly happened in the study? What was the task? How was behavior simulated? etc.

8. p. 11: Can you give one or two concrete examples how SNA may help to inform interventions?

9. p. 11, lines 45-57: This paragraph does not really fit into the flow of the argumentation, as the section overall talks about SNA and not about CSCL. Perhaps it would be good to add a short section "Context of the present study" right before the research questions section that informs the reader about the setting you used (that it was a CSCL setting, what tasks students had etc.). Then the line of argumentation would perhaps be smoother.

10. p. 12: In the research questions section, the disbalance between the rather long description of processes that were uncovered by SNA and the rather short description of the relation between those indicators and student learning in the results section becomes especially obvious: Actually, you do not have a research question for all the results you present between p. 14 to 20! If you take the RQs seriously, you would actually need to label everything that is presented on p. 14 to 20 as "preliminary analyses" that are necessary to finally look at the relation of the indicators and student learning. Thus, again my recommendation would be to shift the length of these parts of the results section towards a stronger balance. E.g., to answer the RQ2, I wonder whether it is really necessary to look at two students' networks to make the case. Perhaps this analysis could be replaced by one that provides more detail to the relations between SNA indicators and student learning outcomes.

11. p. 12: I was wondering how you can empirically answer RQ3. Perhaps it would be wise to take it out as an research question, but talk about its content as part of the discussion section?

12. p. 13: "intended learning outcomes (of the course)" - what were they exactly?

13. p. 13: "Data were cleaned" - what does that mean?

14. p. 14: Can you quickly say what your goals are for each of the three levels of the visualization?

15. p. 14: I am missing information in the results section on how student achievement was measured.
16. p. 16, figures 3A and 3B: I believe there is a problem with the captions of figures 3A and 3B. In the text it says that figure 3A has to do with out-going degree and that it demonstrates the information giving network, while the figure caption says "information-receiving network". Figure 3B should also read "in-going", instead of "out-going", I guess. Regarding figure 3B, I wondered why none of the arrows point to somebody else than the instructor. The figure seems to imply that only the instructor received messages from students, but that virtually never one student received a message from somebody else. This is probably wrong though. I also wondered whether your interpretation that the instructor dominated the discussion is valid. My interpretation of fig 3A and 3B would be that the instructor did receive many messages; but that he actually was not producing so many messages (as indicated by the larger node of S3 and S21 in fig. 3B).

17. p. 17, line 12: It should read "figure 4", not "figure 3").

18. p. 19: The network metrics seem to be interesting, but the information that is conveyed by the values is a bit hard to grasp. E.g., what does a "closeness centrality index" of 0.36 mean? Are there some sort of threshold values by which you could say that closeness centrality is "high" (just as there are conventions regarding for example how to interpret effect sizes or correlations as "weak", "average", or "strong")?

19. p. 21, lines 45-53: formulations such as "degree and out-degree were not significant" are problematic. More precise would be "degree and non-degree did not significantly correlate with student grades. Please check for more of such formulations (e.g. also on p. 27, line 13). Also: the correlation coefficients of degree and out-degree with student grades seem to be very close to significance, and the size of the correlations is also comparable. So I would be more careful with making the implication that in-degree is much more strongly correlated with student achievement than the other two. If you would test whether the different correlations are statistically significantly different from each other, you would probably find that they are not.

20. p. 21, line 50: Nevertheless, I find the result that in-degree centrality is significantly correlated with student grades, while the other two degree indicators are not, theoretically interesting: This seems to imply that it is more important for my learning that I am addressed often by others but rather that I contribute something to the discourse. In the interpretation of this finding, you could refer a bit to recent learning activity models such as the ICAP model by Chi and Wylie (2014).

21. p. 23, line 4: I did not quite get what the advantage of the automated linear regression is compared to a multiple regression. If you use a multiple regression with a backwards procedure, wouldn't you have the same effect as the one you have with the automated linear regression (i.e. in step 1 you would include all SNA predictors, and then you would reduce
them from step to step by the one predictor that explains the least additional evidence)? Also: Is it possible to present all R-squares and p-values for each SNA indicator?

22. p. 23, line 11: "The most important predictors were information centrality 45.2%..." should be changed into something like: "Information centrality explained 45.2 % of the variance in student grades" (if this is true).

23. p. 23, line 14: "The fact that SNA was able to explain..." - I think you cannot say that SNA "is able to do anything". Rather, as said above, it empowers the researcher to do certain things. Similarly on p. 27: The accuracy of the ALM prediction was 37.7%".

24. Finally, although the overall language style is good, there are still a couple of errors across the manuscript. I would therefore encourage the authors to have a native speaker check the revision.

**Are the methods appropriate and well described?**
If not, please specify what is required in your comments to the authors.

No

**Does the work include the necessary controls?**
If not, please specify which controls are required in your comments to the authors.

No

**Are the conclusions drawn adequately supported by the data shown?**
If not, please explain in your comments to the authors.

No

**Are you able to assess any statistics in the manuscript or would you recommend an additional statistical review?**
If an additional statistical review is recommended, please specify what aspects require further assessment in your comments to the editors.

I am able to assess the statistics

**Quality of written English**
Please indicate the quality of language in the manuscript:

Needs some language corrections before being published
Declaration of competing interests
Please complete a declaration of competing interests, considering the following questions:

1. Have you in the past five years received reimbursements, fees, funding, or salary from an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

2. Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

3. Do you hold or are you currently applying for any patents relating to the content of the manuscript?

4. Have you received reimbursements, fees, funding, or salary from an organization that holds or has applied for patents relating to the content of the manuscript?

5. Do you have any other financial competing interests?

6. Do you have any non-financial competing interests in relation to this paper?

If you can answer no to all of the above, write 'I declare that I have no competing interests' below. If your reply is yes to any, please give details below.

I declare that I have no competing interests.

I agree to the open peer review policy of the journal. I understand that my name will be included on my report to the authors and, if the manuscript is accepted for publication, my named report including any attachments I upload will be posted on the website along with the authors' responses. I agree for my report to be made available under an Open Access Creative Commons CC-BY license (http://creativecommons.org/licenses/by/4.0/). I understand that any comments which I do not wish to be included in my named report can be included as confidential comments to the editors, which will not be published.

I agree to the open peer review policy of the journal.