Author’s response to reviews

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

Authors:

Mohammed Saqr (saqr@dsv.su.se)

Uno Fors (uno@dsv.su.se)

Matti Tedre (matti@dsv.su.se)

Version: 2 Date: 20 Oct 2017

Author’s response to reviews:

Dear Editor,

Thank you very much for the positive view of our manuscript, and for the opportunity to revise and improve it. We have tried to fully address all comments from the reviewers and have done the necessary changes that we believe have significantly improved the manuscript.

Technical Comments:

1. Abstract

Please remove the 'Aim of this work' heading from your abstract. This section can be contained within the background section.

That section has been removed and merged with the background section.

2. Consent

Please include an explicit statement regarding consent to participate - please clarify whether students were informed their data would be used for research purposes.

The section on ethics has been updated to reflect the agreement of students that their data would be used in research.
3. Figure legends

Please move the figure legends to the end of your manuscript so that they are clear for our production teams should your manuscript be accepted.

All figure legends were moved to the end of the manuscript, they were also revised and improved.

Reviewer reports:

John Sandars (Reviewer 1): Please include all comments for the authors in this box rather than uploading your report as an attachment. Please only upload as attachments annotated versions of manuscripts, graphs, supporting materials or other aspects of your report which cannot be included in a text format.

Please overwrite this text when adding your comments to the authors.

The authors appear to have appropriately responded to the recommendations made by the previous reviewers.

As noted by the reviewers, the authors present an article that includes both a descriptive practical application of SNA and a more research orientated use of SNA to identify associations with performance. The authors may like to consider a short discussion of the practical use of SNA eg plug-ins.

Thank you very much for the feedback, and for the positive view of our manuscript. The introduction section of SNA has been extended to include some practical uses of SNA (Page 7 1st paragraph), the section about plugins were further extended on page 6, the reference provided has an extensive review of each plugin and how it works.

Robert A. Hanneman (Reviewer 2): The authors have done a fine job in their revision of the article, and have carefully responded to almost all points of the extensive reviews of the earlier submission. The clarification of the focus toward instructional applications, with extensive discussion has greatly clarified and improved the article; it should now be both an interesting research result, and of practical utility to readers.

The analysis and reporting of results, both visually and statistically, are basically sound. Additional analyses of the dynamics and use of multiple outcomes go a long way toward resolving issues of endogeneity, dynamics, and the like.
There is an important remaining statistical issue. Network analysis is not based in independent random samples of observations. The standard errors and significance tests routinely computed by statistical packages assume such independence. The (usually) position network autocorrelation of cases in network samples usually results in the standard errors computed under independence understating variance, and hence overstating statistical significance. For non-independent samples, it is best to use random permutation of the data to construct sample-based standard errors. Other alternatives, not as good, include standard re-sampling estimators.

It would be a good bit of work for the authors to re-do all of the significance testing using permutation methods. It would probably be adequate to add a short discussion of this statistical issue to the paper, urging caution and a conservative approach to interpreting the current significance tests. UCINET might contain a permutation test for the significance of the tau measures - this could be used to assess whether there is a problem (usually, in my experience, there isn't) with the computed standard errors.

Thank you very much for the encouraging review, we really appreciate the positive feedback which had allowed us to improve the manuscript. Based on these recommendations, the correlation tests were re-done with permutation methods, using PAST (Paleontological statistics software package for education and data analysis). The permutation test is based on 9999 random replicates. The results were also double checked using R.

Ingo Kollar (Reviewer 3): The manuscript has substantially improved from the previous to the current version. The authors now clearly demonstrate the potentials that SNA offers when it comes to the analysis of learning processes on different levels (course level, thread level, individual level) and to the explanation of external performance criteria. They also did overall a good job in attributing SNA as a method itself explanatory value and now instead convincingly argue that SNA is one more tool in the researcher's hands to identify processes of collaborative learning that eventually lead to individual performance.

While my overall impression of the manuscript is positive, some minor issues remain that should be addressed in one further revision of the paper:

Thank you very much for the positive feedback and the very detailed review that has allowed us to improve the manuscript a lot, really appreciated.

1. Abstract: "and the relationship of SNA with students' performance remains largely unknown" --> how can a research method have a relation with students' performance? From my point of
view, this sentence should read somehow like this: "and the relationship of parameters that can be obtained via SNA with students' performance remains largely unknown".

Thank you for highlighting this issue, the text was corrected as recommended.

2. Abstract (results): "the interaction patterns and the significant mediators" --> mediators between what variables? Please explain.

This text has been removed and replaced with a clearer phrase. It became “track the knowledge flow, the interaction patterns and identify the active participants and the prominent discussion moderators

3. P. 5, line 37: "...see humans are..." needs to be changed into "...see humans as...

Thank you, this has been fixed as recommended.

4. P. 6: You are critizing dashboards that are usually used in LMSs for their restricted possibility in data analysis. I wondered whether you could give an example of such an LMS and the analytic dashboard it uses. Perhaps you could even present a screenshot that makes its restrictions visible?

Learning management systems have very limited analytics dashboards, that has been recognized by the major manufacturers, for example, Blackboard has acquired a company called X-ray analytics and it is still very limited and has not offered social network analysis; Moodle launched project Inspire to offer the functionalities, the project has no plans to deliver SNA in the near future according to personal communication with the team leader, external plugins seems to work with specific versions only, so plugins like SNAPP and Graphfes don’t work with current versions of LMSs.

We have fixed this and offered examples of the dashboards of the major LMSs and explained their capabilities, the reference of this section has an excellent review of the capabilities of external add-ons commonly used to generate social network analysis with screenshots of each application. Page 6 3rd paragraph

5. Maybe my most important remark: I still think that the different SNA parameters you presented could be better introduced. I think the reader would benefit much from an explanation of how each of the parameters are computed when these explanations are already
presented in the theoretical section, i.e. on p. 7 and 8. E.g. at the bottom half of p.8, you say what "centrality" means, but it remains unclear how it is computed. The same applies for the introduction of all other parameters in that section ("closeness centrality", "weighted degree centrality" etc.).

Thank you for helping improving this section. We have amended these centrality measures on page 9 with explanations for each one, now we think it is easier for the reader to get what they mean, later in the methods section, there is a comprehensive overview of each parameter, that was also improved for each parameter and now with explanation of how it is computed.

6. Also on p.7, you say that SNA offers you to delineate the "importance" in mediating interactions. Here I wondered what you exactly mean by "importance". At first sight, I would understand "importance" (e.g. of a contribution) as something that requires content-analytical techniques to assess. You however seem to mean something else (I assume the importance of a person as indicated by in- and out-degree centrality). Please explain in a clearer way what you mean here.

Thank you, we have tried to take away the ambiguity form this sentence and written it in a much clearer way. See for example the figure (1) legend, now moved to the end of manuscript page 37 as recommended by the editor.

7. I very much appreciate that the authors ask clear research questions. Yet, the results section could do a better job in describing what analysis refers to what research question. This could be done by using appropriate headings such as "The social structure and performance (research question 2)" and by possibly repeating each research question in the results section before the respective analyses are presented.

Thank you, we have added research questions as headings at the appropriate sections as recommended.

8. Research question 3 seems to be broader than the analyses you ran. It reads "How can student's position, interactions, and relations in a network be used to predict his or her final performance?". In the analysis, you however restrict that question to the identification of underachievers (which I - by the way - do not understand how you identified them -should be described as well). I would thus recommend to be more specific in the formulation of the research question, i.e. to restrict it to the identification of underachievers.
There is a section in the method that describe how we measure performance and what we mean by underachievers, in the second paragraph of page 12

"Performance: Students’ final grades were used to measure final course performance, and midterm grades were used to measure performance up to the midterm. Objective Structured Clinical Exam grades were used to measure clinical performance, and multiple-choice questions (MCQs) were used to measure knowledge comprehension, analysis, and application. Students were classified either as underachievers who are at risk of failing (lowest 1/3) or achievers who are relatively safe from failing (top 2/3)."

There are two sections in this research question, the first is predicting performance, for which we presented Table (3) for predicting performance using SNA predictors calculated at the end of course and it tries to predict performance using three measures according to reviewer previous recommendations, these were Final Grade, Clinical results and MCQ.

And Table (4) for predicting performance using SNA predictors that were calculated at midterm, which tried to use the early participation parameters to predict both Final Grade and Midterm grade. So, there are five tables that justifies the “predicting performance” research question part. As recommended in the previous research question, we have added the research question appropriately before that section so it is clearer to the reader.

The prediction of under achievers is the second part of this research question, and it is answered by the table (5).

We think that appropriately placing the research question have clarified this point now.

9. P. 13: I do not quite understand why "role in moderating and relay of information" is bolded and why it is presented before the parameter (as this is always done the other way around in case of the other parameters that are presented in bullet points".

Thank you for this comment, Yes, we have fixed this and it matches the general formatting now and is presented as a heading as the other sections of parameters.

10. P. 13: I am still not completely sure if I understood correctly what you mean by betweenness centrality (does this mean the frequency by which a person has received information from one participant and passed it on to other participants?).

The between centrality were explained in more details and the way it was computed was also explained. Page 14
Please note that other reviewers have commented on the detailed information about centrality measures citing the availability of these information from different sources, according to the second reviewer comments:

“I would suggest that much of the material explaining specific network concepts (e.g. the many variations of centrality measures) be simply dropped - relying on available good reference works.”

11. P. 13: Could you add information on how information centrality and closeness centrality are actually computed (right now, you rather abstractly talk about those two concepts). The same goes for "clustering coefficient".

The information centrality, closeness centrality (Page 14) and Clustering coefficient (Page 16) were explained and the way they are computed was also clarified.

12. P. 14 (bottom): please add information on what research questions are addressed by (a) network visualization and (b) network analysis.

Thank you, as was recommended in a previous comment, we have placed the research questions in the appropriate section, so the sections are more understandable now.

13. Figure 3b: I am still not sure if I understand the "out-degree centrality" figure. What exactly is the difference to figure 3a? From my understanding, 3a shows that the teacher received a lot of messages from many students; and in 3b, I take that students give a lot of information to the teacher. So, what is the difference between the two visualizations, also given the fact that they look very similar to each other?

While the teacher position and the information he received might be similar, Students are different; figure 3B highlights the students who received information and thus they were prominent information givers. In a collaborative environment, one would expect more students to receive information as their peers value the information they give and interact with it.

14. P. 20: Again, please add what research question you tackle under the heading "network properties and correlation to performance".

Thank you, as was recommended in a previous comment, we have placed the research questions in the appropriate section, so these sections are more understandable now.
15. P. 20: perhaps I missed it - but if not: please explain what you mean by "graph density", "average degree" and "average clustering coefficient".

There was an explanation of these concepts, however, as we expanded explanation of other indicators, we detailed these concepts too with improved clarity. On page 13

16. P. 22: Again, please add what research question you tackle under the heading "Can social network analysis predict performance" (and as recommended above: please check whether this question could be formulated in a more specific way -- underachievers).

Thank you, as was recommended in a previous comment, we have placed the research questions in the appropriate section, so these sections are more understandable now.

17. As I am not familiar with ALM, I was wondering whether its results should include more information than just the amount of explained variance. What about beta-weights and p-values in tables 3 and 4? At least for "regular" regression, I would expect this information to be presented.

Thank you for raising this issue, in this study, we offer two predictive models. The first type is predictive modelling, creating a predictive model would demonstrate the possibility of using SNA data to enhance our understanding of the interactions in the learning environment, and thus enable us to improve it, as well as to predict low grades, so that low achievers in next courses can be supported. The predictive model we needed should demonstrate the capabilities of normalizing outliers, include all predictors, detecting and removing similar predictors; besides, it needs to be automatic to demonstrate the feasibility of using it in a learning analytic dashboard that offers insights with no human intervention with the data. Since simple linear regression lacks these features, we chose automatic linear regression, the model has all the previous features with ensemble capabilities, and we used the ensemble method bootstrap aggregation (Bagging). Bagging is a machine learning algorithm that is designed to enhance accuracy and stability of the model with reduction of variance and overfitting.

We ran the model 100 times, each time, the models selects a different random subsample, the result is the averaging of repetitive models. SPSS reports of ALM regression results contain the R2 and the significant predictors and how much they contributed to the model or (“the residual sum of squares if the predictor was excluded from the model. The values are normalized so that the sum is 100 %”). The use of automated model, although might not be favorable by some statisticians, for two main reasons; the first, to show the possibility of offering automated prediction which is what Learning Analytics is about, Second to rank predictors as a method to
choose the most important predictors for stepwise regression. Acknowledging the shortcoming of this method, we have also included a standard backward multiple regression (based on a previous recommendation of the reviewer), the second model (explanatory model) uses causal explanation by ways of statistical modeling to build and test a hypothesis. The second model was stepwise backward binary logistic regression, predictors included age, previous performance and SNA predictors and reported regression performance tests such as chi-square test of independence, Cox & Snell R Square, Nagelkerke R Square, Hosmer and Lemeshow test.

18. Table 5: Maybe I missed it, but did you explain how you identified students as "at-risk" or "safe"? Based on the grades in the final exams?

There is a section of the method that describe how we measure performance and what we mean by underachievers. This is indicated in the second paragraph of page 12

Performance: Students’ final grades were used to measure final course performance, and midterm grades were used to measure performance up to the midterm. Objective Structured Clinical Exam grades were used to measure clinical performance, and multiple-choice questions (MCQs) were used to measure knowledge comprehension, analysis, and application. Students were classified either as underachievers who are at risk of failing (lowest 1/3) or achievers who are relatively safe from failing (top 2/3).

19. P. 27, second paragraph: Here you start the discussion saying "On the course level..." - I was subsequently missing sentences that would start with "On the thread level" and "on the individual level", i.e. you could perhaps go through these three levels in a more systematic way.

We have added these sentences in the appropriate places, as recommend.

20. Although the English is very good, especially the newly added or revised parts need some additional language editing (e.g., I see articles such as "a" or "the" missing here and there).

Thank you, a thorough revision of all the manuscript was done and these were corrected.