Fostering a Learning Community in MOOCs

Cindy E. Hmelo-Silver	Carolyn P. Rosé	Jeff Levy
Indiana University	Carnegie Mellon	OfficeHours
543 Eigenmann Hall	University5000 Forbes Avenue	817 W. Peachtree St
Bloomington IN 47406	Pittsburgh, PA 15217	Atlanta GA 30308

1. Introduction

A key hurdle preventing MOOCs from reaching their transformative potential is that they fail to provide a social environment that is conducive to sustained collaborative engagement and learning. Correlational analyses from existing MOOCs demonstrate a reliable connection between social integration into the threaded discussions and course retention. The best evidence so far suggests an important connection between social support and retention consistent with findings in other types of online communities. If we can engineer MOOCs to achieve greater success at providing affordances for sustained social engagement and learning, the potential for impact is immense.

2. Facilitating Small Groups

Introducing facilitation into large scale threaded discussion in MOOCs requires extensions to existing wisdom on facilitation of small groups [1; 2]. We seek to understand how better to facilitate student-centered learning and student agency (and deep engagement) on a grander scale in online environments. We know a lot about strategies that can be used to support students in collaborative projects and problem based learning [3] but these have generally only been used on a small scale because they requires intense monitoring of the collaborative discourse for both collaboration and content (what we might term productive collaboration). Engineering facilitation requires attention at two levels: supporting the collaboration among learners, but also supporting facilitation for teaching assistants or other tutors.

To accomplish larger scale facilitation, there are two important aspects to consider. One is identify markers of both collaborative activities and disciplinary content and/or practices. Some of these might include indicators such as making contributions, asking responding to questions, social network analysis, or indicators of transactivity. Learning analytics (LA) techniques are important to be able to accomplish this [4; 5]. The markers of disciplinary content and practices may be more difficult to measure. The second aspect is acting on what is learned through LA. A first step might be a dashboard that might alert instructors as to where they might need to intervene as well as suggesting appropriate strategies.

Research on these massive networked communities should build on research on best practices in CSCL and problem-based learning (PBL; e.g., [3; 6-10]). This research suggests that certain instructional goals can support productive collaboration, and that there are particular cognitive, metacognitive, and social strategies that can best serve particular kinds of goals. For example, in some groups such as a PBL group, goals might include:

Educational Goals for students:

- E1. Construct causal explanations
- E2. Employ effective reasoning and argumentation.
- E3. Identify knowledge limitations
- E4. Self-directed study
- E5. Evaluate their learning and performance.

Performance goals for facilitator

- P1. Keep all students active in the learning process
- P2. Keep the learning process on track
- P3. Make the student's thoughts and their depth of understanding apparent
- P4. Encourage student reliance on selves and peers for direction and information.

Strategies that can support these goals are shown in Table 1.

Table 1. Example Strategies (adapted from [3])

Strategy	Goals	How goals accomplished?
Use of open-ended and metacognitive questioning ("Does everyone agree with Cindy"; What do you mean by that)?	E1-4, P1, P3, P4	General strategy to encourage explanations and recognition of knowledge limitations
Pushing for explanation (e.g., asking "Why")	E1, P3 E3, P3, P4	Construct causal models Students realize limits of their knowledge
Revoicing	E1, P2 P1 P2	Clarify ideas Legitimate ideas of low status students Mark ideas as important and subtly influence direction of discussion
Summarizing	E4, P1 P1 E1, E5, P3 P2 E5, P3	Ensure joint representation of problem Involve less vocal students Help students synthesize data Move group along in process Reveals facts that students think are important
Generate/ evaluate hypotheses (initiate brainstorming or ask group where we are now with the hypotheses identified thus far).	E2, E4, P2 E1, E2, P3, P4	Help students focus their inquiry Examine fit between hypotheses and accumulating evidence
Encourage construction of visual representation (e.g., flow charts, table, concept map)	E1, E5, P3	Construct integrated knowledge structure that ties mechanisms to observable effects

For TA's or any kind of automated agents to use these strategies, it will be important to have indicators so that whatever the facilitation strategy needed, there is some way to first indicate that attention is needed in a group. Some of this might be timed--for example, having a group summarize where they are could be done at intervals based on the ebb and flow of course assignments and content difficulty. Some of the social prompts could be based on shallow indicators like participation and length of posts. Others would need more semantically oriented indicators. For example, the first time that someone mentions or proposes some complex idea, it would be reasonable to ask the group if they understand. Semantic indicators could compare what is being discussed with key course objectives and past discussion. To the extent that these indicators can be identified, these are other places where simple intervention might be quite productive. An area ripe for research is identifying these indicators, understanding what kinds of strategies might be generally useful and the extent to which these indicators are amenable to computational solutions.

In our computational work, we have been developing techniques for extracting indicators within messages that tell us something about a student's orientation at a time point (i.e., a week of participation within the course) towards the course, which will enable us to make a prediction about how likely it is that the student will drop out of the course on the next time point. The relevant levels of analysis are post, week, and trajectory within the course. First, we extract indicators from individual messages. Then we aggregate messages within a week in order to construct an indicator for a week. We then use a survival analysis to model the probability that the student with drop out on the next time point given the value of the variable. In this way, we can identify risk factors that instructors should be aware of so that they can focus their efforts to support those students. Indicators we have explored include social network analysis measures such as authority and hub scores [2], indicators of social subgraph membership [1] measures of motivation and cognitive engagement [11], and indicators of relationship formation and measures of loss of relationship due to attrition of other students [12]. Overall, we find that measures of high authority and hub scores as well as high numbers of formed relationships predict lower attrition, as do measures of high motivation and high cognitive engagement. Conversely, measures of relationship loss predict higher dropout. Similarly, membership in subgroups where there is high attrition also predicts higher attrition. To model relationship formation over time, we are experimenting with probabilistic graphical models that combine text, social network, and thread structure representation in order to identify coordinated group behavior that may indicate the formation of a subcommunity [13]. Beyond measures at the student level that pick out students who are at risk and may need extra attention, we use structural equation modeling techniques to identify the factors that affect whether a thread that is started by a student who is reaching out for help will get a satisfactory response. In order to reduce the load on instructors, we are developing matrix factorization techniques to identify community members whose effort we may be able to enlist to respond to at risk threads [14].

Recent research in the learning sciences has begun to address how to use LA productively [15]. At the same time, MOOCs can provide transformative opportunities for learners if we can identify and support learning communities within the larger communities. Bringing the LA and learning sciences communities together for discussion and joint efforts can provide opportunities to better understand how to facilitate learning communities in this frontier.

3. References

[1] Rosé, C.P., Carlson, R., Yang, D., Wen, M., Resnick, L., Goldman, P., and Sherer, J., 2014. Social Factors that Contribute to Attrition in MOOCs. In *Proceedings of the First ACM Conference on Learning @ Scale*.

[2] Yang, D., Sinha, T., Adamson, D., and Rosé, C.P., 2013. Turn on, Tune in, Drop out: Anticipating student dropouts in Massive Open Online Courses. In *Proceedings of the NIPS Data-Driven Education Workshop* (2013).

[3] Hmelo-Silver, C.E. and Barrows, H.S., 2006. Goals and strategies of a problem-based learning facilitator. *Interdisciplinary Journal of Problem-based Learning 1*, 21-39.

[4] Mu, J., Stegmann, K., Mayfield, E., Rosé, C.P., and Fischer, F., 2012. The ACODEA Framework: Developing Segmentation and Classification Schemes for Fully Automatic Analysis of Online Discussions. *International Journal of Computer Supported Collaborative Learning 7*, 285-305.

[5] Gweon, G., Jain, M., Mc Donough, J., Raj, B., and Rosé, C.P., 2013. Measuring Prevalence of Other-Oriented Transactive Contributions Using an Automated Measure of Speech Style Accommodation. *International Journal of Computer Supported Collaborative Learning 8*, 245-265.

[6] Bielaczyc, K., Kapur, M., and Collins, A., 2013. Cultivating a community of learners in a K-12 classroom. In *International Handbook of Collaborative Learning*, C.E. Hmelo-Silver, C.A. Chinn, C.K.K. Chan and A.M. O'Donnell Eds. Routledge, New York, 233-249.

[7] Chan, C.K.K., 2013. Collaborative knowledge building: Towards a knowledge creation perspective. In *International Handbook of Collaborative Learning*, C.E. Hmelo-Silver, C.A. Chinn, C.K.K. Chan and A.M. O'Donnell Eds. Taylor and Francis, New York, 437-461.

[8] Chinn, C.A. and Clark, D.B., 2013. Learning through collaborative argumentation. In *International Handbook of Collaborative Learning*, C. Hmelo-Silver, C.A. Chinn, C.K.K. Chan and A.M. O'Donnell Eds. Routledge, New York, 314-332.

[9] Fischer, F., Kollar, I., Weinberger, A., Stegmann, K., Wecker, C., and Zottmann, J., 2013. Collaboration Scripts in Computer-Supported Collaborative Learning. *International Handbook of Collaborative Learning, Routledge, New York*, 403-419.

[10] Schwarz, B.B. and Asterhan, C.S., 2011. E-Moderation of synchronous discussions in educational settings: A nascent practice. *Journal of the Learning Sciences 20*.

[11] Wen, M., Yang, D., and Rosé, C.P., under review. Linguistic Reflections of Student Engagement in Massive Open Online Courses. submitted to *International Conference on Weblogs and Social Media*.

[12] Yang, D., Wen, M., and Rosé, C.P., under review-a Peer Influence on Student Attrition in Massively Open Online Courses. *submitted to Educational Data Mining.*

[13] Kumar, A., Palakodety, S., Wang, C., Wen, M., Rosé, C.P., and Xing, E., under review. Large scale structure aware community discovery in online forums. *submitted to Knowledge Discovery and Data Mining*.

[14] Yang, D., Wen, M., and Rosé, C.P., under review-b. Towards Increasing the Resolvability of Unresolved Threads in MOOCs. submitted to *International Conference on Weblogs and Social Media*.

[15] Martin, T. and Sherin, B., 2013. Learning analytics and computational techniques for detecting and evaluating patterns in learning: An introduction to the special issue. *Journal of the Learning Sciences 22*, 511-520.