From Insights to Interventions: Informed Design of Discussion Affordances for Natural Collaborative Exchange

Sreecharan Sankaranarayanan, Gaurav Singh Tomar, Miaomiao Wen, Akash Bharadwaj and Carolyn Penstein Rosé
{sreechas, gtomar, mwen, akashb, cprose}@cs.cmu.edu
Language Technologies Institute
Carnegie Mellon University, Pittsburgh, PA, USA

Abstract

Despite studies showing collaboration to be beneficial both in terms of student satisfaction and learning, isolation is the norm in MOOCs. Two problems limiting the success of collaboration in MOOCs are the lack of support for team formation and structured collaboration support. Lack of support and strategies for team formation prevents teams from being set up for success from the beginning. Lack of structured support during synchronous collaboration has been demonstrated to produce significantly less learning than supported collaboration. This paper describes a deliberation based team formation approach and a scripted collaboration framework for MOOCs aimed at addressing these problems under the umbrella of DANCE [1] whose overarching focus is the enhancement of team-based MOOCs. These two examples of current work have been used as illustrations of insights informing interventions in MOOCs.

Introduction

Past work and analysis of data from Massive Open Online Courses (MOOCs) has revealed evidence of the association between types of conversational interactions and retention (Wen, Yang, and Rose 2014a; 2014b; Wen, Yang, and Rosé 2015), team project quality (Yang, Wen, and Rosé 2015) and learning (Wang et al. 2015). These insights have ultimately informed the design of interventions that support improved outcomes (Ferschke et al. 2015a; 2015b; Howley et al. 2015). This paper summarizes findings from a series of investigations related to the design and development of such interventions under the umbrella partnership with edX, referred to as Discussion Affordances for Natural Collaborative Exchange (DANCE), with a focus on enhancing team-based MOOCs.

Our work began with two interventions developed for the edX Data, Analytics, and Learning (DALMOOC) course (with lead instructor George Siemens from UT Arlington and co-instructors Carolyn Rosé from Carnegie Mellon University, Dragan Gasevic from the University of Edinburgh and Ryan Baker from Teachers College), which was launched in early October of 2014 (Rosé et al. 2015). The first of these interventions built on a social recommendation algorithm that considered constraints both on the part of the person being recommended as well as the person the recommendation was addressed to (Yang, Adamson, and Rosé 2014). This algorithm was built into a MOOC intervention called Quick Helper, which was an augmentation of the edX threaded discussion forums (Howley et al. 2015) that recommended other participants with expertise relevant to the question being asked. Participants could then choose to ask these recommended users to answer the questions, thereby effectively addressing help-seeking in MOOCs. A second intervention was a collaborative reflection chat intervention called Bazaar, building on earlier work in classroom contexts (Adamson et al. 2014). In the MOOC setting, it was introduced as a collaborative component to be completed by students in each unit after having done the individual learning for that unit (Ferschke et al. 2015a; 2015b). In DALMOOC, use of Quick Helper increasingly overtook the standard approach to participation in threaded discussions throughout the course, and participation in the synchronous collaborative chats was associated with substantial local reductions in probability of drop out (Ferschke et al. 2015b).

Both of these initial interventions were designed for short, periodic collaborative exchanges. More recently we have been working towards more persistent social interaction throughout a course in the form of team based projects. Our analysis of data from two team based MOOCs suggests that the success of teams in state-of-the-art team based MOOCs is low (Wen, Yang, and Rosé 2015; Yang, Wen, and Rosé 2015). While the behavior of team leaders, and to a lesser extent that of other team members, predict team outcomes, the evidence points to the conclusion that the problem starts even before the teams begin to function in that capacity. In particular, the team formation process itself must be improved in order to produce teams that are positioned for success from the start.

In previous work, team formation has occurred early in the course through personal messaging interactions based on perceptions around learner profiles like prior knowledge, language, location and other demographics. We developed a design for a deliberation based team formation procedure, which we worked to refine and validate in pilot experiments conducted in the crowdsourcing environment, Amazon Me-

[1] Discussion Affordances for Natural Collaborative Exchange - dance.cs.cmu.edu

chanical Turk. That procedure offered opportunities for students to interact in a project-relevant way in the discussion forums with members of the whole course community prior to team formation. Based on observed discussion processes, which can be automatically analyzed, teams were formed using a constraint satisfaction algorithm that is able to assign all students to teams in such a way that the average team quality is maximized (by a criterion based on qualities of observed interaction in the discussion forum).

In our continued work we are using the insights gained from these studies to form design principles with high internal validity (using controlled experimentation within the crowdsourcing environment) which we will then test in realistic deployment on the edX platform. The design principles are being encoded into future offerings of DANCE as a way of demonstrating the integration and study of artificial intelligence in MOOC settings.

The rest of the paper will first discuss the resources currently available through DANCE as well as planned additions to the DANCE repertoire for supporting scripted collaboration in MOOCs. The section on "Deliberation-based Team Formation" describes how insights from a crowdsourced study are being replicated in a real MOOC deployment so that intelligent support for team formation can be incorporated to set teams up for success.

Discussion Affordances for Natural Collaborative Exchange

DANCE is a satellite working group that seeks to draw from two decades of research in Computer Supported Collaborative Learning to enhance instructionally beneficial discussion opportunities available to students involving researchers and developers from multiple universities, foundations and industrial organizations. A particular goal is to apply artificial intelligence approaches, such as text mining, recommender systems, and intelligent conversational agents to make online collaborative learning experiences effective.

Another goal of DANCE is to aggregate CSCL/MOOC resources in the discussion affordances space and make them available to the community at large. A Google Scholar account[2] and a resources page[3] serve this end. DANCE will offer DiscourseDB[4], a data infrastructure project designed to bridge data sources from multiple platforms for hosting different learning experiences with a vision of providing a common data model designed to accommodate data from diverse sources including but not limited to chat, threaded discussions, blogs, twitter, wikis, and text messages. A discussion forum XBlock for the edX platform will also be released as a first step toward enabling script based support for collaboration in MOOCs.

Scripted Collaboration in MOOCs

Despite studies indicating the advantages of scripted collaboration with respect to learning and student satisfaction among an older student demographic (Ke and Xie 2009; Keeton 2004) akin to those found in MOOCs, isolation has been the norm in MOOCs, and so far there has been a shortage of affordances for collaborative learning experiences in those platforms. One notable exception is the NovoEd platform[5] which frequently requires students to work in groups to successfully complete courses. The "Technology Entrepreneurship" course offered on the platform reported a completion rate of 65%(Eesley 2014) which seems remarkable in comparison with the far lower completion rates reported by standard MOOCs. This suggests other platforms might benefit from offering collaboration opportunities for students, especially if they can be properly structured and supported.

Encouraged by this, DANCE is making progress towards supporting scripted collaboration in MOOCs using two major approaches.

- **Learning Tools Interoperability (LTI)** - LTI defines the integration of third party tools into LTI compliant MOOC platforms such as edX and Coursera. Bazaar, the collaborative reflection chat intervention has been successfully integrated into edX course using this method.
- **Open edX XBlocks** - An XBlock is a modular and hierarchical construct used in the Open edX. An XBlock is analogous to a div in a web application.

The LTI approach is appealing because it enables one time development and multiple deployments. It is however reliant on the platform to expose services and data in an LTI compliant fashion. The standard itself is amorphous and may change with time and requirements. The XBlock on the other hand is tightly coupled to the edX platform. edX's own use of the XBlock framework guarantees availability and compliance. It also allows us to anticipate and plan integrations with features like Teams[6] and Cohorts[7] which are scheduled to be released soon. The XBlock can potentially be used by all platforms that run on the open edX framework including Lagunita[8], MOOC.org[9] and Google Course Builder. An overview of the kinds of support that the XBlock enables has been described below.

Our XBlock design is set up to be customizable, providing opportunities for individual learning, potentially supported by intelligent tutoring systems augmented with collaborative phases. Thus, the XBlock is customizable, with a configurable pre-collaboration step that is an individual activity that provides instructions and preparation materials. Materials in the form of audio, text and video can be embedded using edX Studio without the need to modify the XBlock code. Group matching and role allocation by default is random.

Due to the varied nature of an intelligent tutoring system (ITS), the functionality isn't planned on being built into the XBlock. The ITS can instead be integrated by hosting it as

[2]dance.cs.cmu.edu/scholar
[3]dance.cs.cmu.edu/resources
[4]discoursedb.github.io
[5]novoed.com
[6]open.edx.org/features/teams
[7]open.edx.org/features/cohort-specific-content
[8]lagunita.stanford.edu/about
[9]www.mooc.org

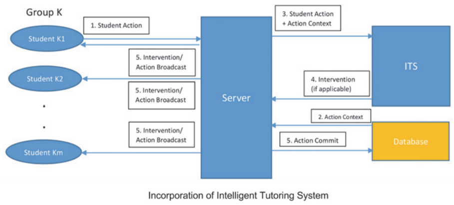

Figure 1: Incorporation of an Intelligent Tutoring System

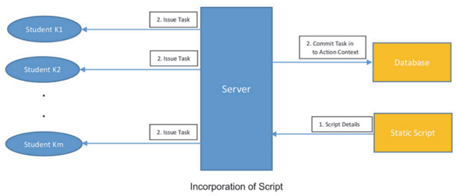

Figure 2: Incroporation of a Script

an external service or by modifying the XBlock implementation. The incorporation process is shown in Figure 1.

For the collaboration phase, simple automated scripts can be natively implemented in the XBlock in the absence of the above ITS design. The server may have to issue tasks based on a static script as shown in Figure 2.

Example: Deliberation-based Team Formation

Few online communities offer support for team formation other than self selection and without careful orchestration many teams have been found to fail (Kraut and Fiore 2014).

We therefore propose a team formation strategy that allows participants to observe and participate in interaction on the discussion forum as preparation for collaboration. Given also that processes like transactive reasoning, knowledge convergence and common ground underpin a successful collaboration (Teasley et al. 2008), we use automatic transactivity identification (Rosé et al. 2008) to group pairs of students who already display successful team processes with each other. In this work we seek to answer the following two research questions.

- Will exposure to community discussions lead to more successful small group collaboration?
  Answered by studying teams formed before the task in comparison with teams formed after deliberation.
- Can evidence of transactive discussions during deliberation inform the formation of more successful teams?
  Answered by studying teams formed randomly after deliberation versus teams formed based on evidence of successful collaborative processes observed during deliberation.

In preparation for an intervention in a real MOOC the following MTurk study was conducted.

Experimental Setup

The collaboration involved students working in groups to accomplish a knowledge integration task. The collaborative task required participants to justify the selection of one of four given energy plans for a city. An energy plan is a combination of up to 4 different energy sources, each of which is represented by one of the group participants as part of a Jigsaw paradigm. In the constraint satisfaction proposal writing task then, the goal was to require each member to represent their assigned perspective also considering how it relates to the perspectives of others in the group. A four-step experimental procedure was designed as follows.

- **Preparation** - To prepare for the Jigsaw task, each participant was assigned to read an article about the pros and cons of a single energy source. To reinforce their learning, a multiple choice quiz with feedback was required to be answered.
- **Pre-task** - Given a city with 5 requirements, the individual was required to write a proposal, which was then posted to the discussion forum.
- **Deliberation** - Each participant was then required to write at least 5 replies to the other proposals. To encourage transactivity, the task instruction when replying to a post encouraged participants to ”elaborate, build upon, question or argue against the ideas presented in that post, drawing from the argumentation in your own proposal where appropriate.”
- **Collaboration** - The participants were then required to collaboratively write a proposal recommending one of the four suggested energy plans for the city. Participants were also required to fill out a survey measuring perceived group outcomes after collaboration.

Experiment 1 - Group Transition Timing

A before-deliberation and after-deliberation condition was set up for comparison. For the after-deliberation condition, in the deliberation step, participants could potentially interact with other participants both inside and outside their group. In the before-deliberation condition, each team was given a separate forum in which to interact with their teammates.

The before-deliberation condition is similar to the current online contests and team-based MOOCs where teams are formed early in the contest or course and then only interact with their teammates after team formation. By comparing these two conditions, we test our hypothesis that exposure to deliberation within a larger community prior to team formation will improve group performance.

Results show that teams exposed to community deliberation prior to group work demonstrate better team performance and that the individuals participants also demonstrated higher-quality individual contributions as well.

Experiment 2 - Grouping Criteria

This time teams in both conditions were grouped after experiencing the deliberation step in the community context. In the experimental condition, which we termed the Transactivity Maximization condition, we applied a constraint satisfaction paradigm that preferred to maximize the extent to which workers assigned to the same team had participated in transactive exchanges in the deliberation with the other people assigned to the same team. In the control condition, which we termed the Random condition, teams were formed by random assignment. In both conditions, the same Jigsaw paradigm applied in the first study was similarly applied.

Results show that teams assigned based on observed transactive communication during deliberation demonstrate better team performance. There was a significant effect on the length of discussion, showing that there was more discussion around the collaborative task in the Transactivity Maximization condition as well.

Design of the Intervention

Given the results of the MTurk study, we claim that teams with the benefit of exposure to the community during a pre-teamwork deliberation process will demonstrate advantages in terms of team performance and teams that experience more transactive communication during a pre-teamwork deliberation process will demonstrate advantages in terms of team performance.

By learning high internal validity design principles from this study, we will be implementing a deliberation-based group formation intervention in a future offering of the MOOC "The Rise of Superheroes and Their Impact On Pop Culture" on edX. The experiment on the course is now designed as follows following the progression of the course flow.

- **Preparation** - All students receive basic instruction about the genre of a superhero story and the eras of superheroes. Students choose either a creative track or a historical emphasis track, which influences how they focus their coursework starting in the first week. (Week 1)
- **Pre-task + Deliberation** - Students start to make concrete their ideas of a superhero as they discuss that concept in connection with a current event from the newspaper. They post this to the discussion forum to discuss and get feedback from other coursemates, after which they are assigned to team in such a way that observed Transactivity is maximized (as in the MTurk study) and each team has two students from the creative track and 2 from the history emphasis track (Week 3).
- **Collaboration** - In the final week of the class, teams then collaboratively develop a time-travel story line where superheroes selected or developed by the students in the team find themselves in a new era. This provides an ideal opportunity for students working on different time periods to collaboratively develop one cohesive storyline. (Week 5 and 6)

Concluding Remarks

In this paper we have briefly offered and overview of the overarching goals of the DANCE initiative, resources that are on offered, and a few concrete examples of script-based collaboration in MOOCs, including deliberation-based team formation. This overview illustrated how insights from studies can be used to design intelligent and automated support for group learning in MOOCs. One of the objectives of the initiative is to bring developers and researchers from various organizations and universities together to "collaboratively" work on the development of collaborative affordances. Collaborators are welcome to get in touch with us.

References

Adamson, D.; Dyke, G.; Jang, H.; and Rosé, C. P. 2014. Towards an agile approach to adapting dynamic collaboration support to student needs. *International Journal of Artificial Intelligence in Education* 24(1):92–124.

Eesley, C. 2014. Virtual courses that avert online education's huge dropout rate.

Ferschke, O.; Howley, I.; Tomar, G.; and Yang, D. 2015a. Fostering discussion across communication media in massive open online courses.

Ferschke, O.; Yang, D.; Tomar, G.; and Rosé, C. P. 2015b. Positive impact of collaborative chat participation in an edx mooc. In *Artificial Intelligence in Education*, 115–124. Springer.

Howley, I.; Tomar, G.; Yang, D.; Ferschke, O.; and Rosé, C. P. 2015. Expectancy value theory of help seeking applied to features in moocs.

Ke, F., and Xie, K. 2009. Toward deep learning for adult students in online courses. *The Internet and Higher Education* 12(3):136–145.

Keeton, M. T. 2004. Best online instructional practices: Report of phase i of an ongoing study. *Journal of Asynchronous Learning Networks* 8(2):75–100.

Kraut, R. E., and Fiore, A. T. 2014. The role of founders in building online groups. In *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing*, 722–732. ACM.

Rosé, C.; Wang, Y.-C.; Cui, Y.; Arguello, J.; Stegmann, K.; Weinberger, A.; and Fischer, F. 2008. Analyzing collaborative learning processes automatically: Exploiting the advances of computational linguistics in computer-supported collaborative learning. *International journal of computer-supported collaborative learning* 3(3):237–271.

Rosé, C. P.; Ferschke, O.; Tomar, G.; Yang, D.; Howley, I.; Aleven, V.; Siemens, G.; Crosslin, M.; and Gasevic, D. 2015. Challenges and opportunities of dual-layer moocs: Reflections from an edx deployment study.

Teasley, S. D.; Fischer, F.; Weinberger, A.; Stegmann, K.; Dillenbourg, P.; Kapur, M.; and Chi, M. 2008. Cognitive convergence in collaborative learning. In *Proceedings of the 8th international conference on International conference for the learning sciences-Volume 3*, 360–367. International Society of the Learning Sciences.

Wang, X.; Yang, D.; Wen, M.; Koedinger, K.; and Rosé, C. P. 2015. Investigating how students cognitive behavior in mooc discussion forums affect learning gains.

Wen, M.; Yang, D.; and Rose, C. 2014a. Sentiment analysis in mooc discussion forums: What does it tell us? In *Educational Data Mining 2014*.

Wen, M.; Yang, D.; and Rose, C. P. 2014b. Linguistic reflections of student engagement in massive open online courses. In *Proceedings of the International Conference on Weblogs and Social Media*.

Wen, M.; Yang, D.; and Rosé, C. P. 2015. Virtual teams in massive open online courses. In *Artificial Intelligence in Education*, 820–824. Springer.

Yang, D.; Adamson, D.; and Rosé, C. P. 2014. Question recommendation with constraints for massive open online courses. In *Proceedings of the 8th ACM Conference on Recommender systems*, 49–56. ACM.

Yang, D.; Wen, M.; and Rosé, C. P. 2015. Weakly supervised role identification in teamwork interactions. In *Proceedings of the 53rd Annual Meeting of the Association for Computional Linguistics*.