

August 30, 2016  
DRAFT

**Investigating Virtual Teams in Massive Open  
Online Courses: Deliberation-based Virtual  
Team Formation, Discussion Mining and  
Support**

Miaomiao Wen

September 2016

School of Computer Science  
Carnegie Mellon University  
Pittsburgh, PA 15213

**Thesis Committee:**

Carolyn P. Rosé (Carnegie Mellon University), Chair  
James Herbsleb (Carnegie Mellon University)  
Steven Dow (University of California, San Diego)  
Anne Trumbore (Wharton Online Learning Initiatives)  
Candace Thille (Stanford University)

*Submitted in partial fulfillment of the requirements  
for the degree of Doctor of Philosophy.*

August 30, 2016  
DRAFT

**Keywords:** Massive Open Online Course, MOOC, virtual team, transactivity, online learning

August 30, 2016  
DRAFT

**Abstract**

To help learners foster the competencies of collaboration and communication in practice, there has been interest in incorporating a collaborative team-based learning component in Massive Open Online Courses (MOOCs) ever since the beginning. Most researchers agree that simply placing students in small groups does not guarantee that learning will occur. In previous work, team formation in MOOCs occurs through personal messaging early in a course and is typically based on scant learner profiles, e.g. demographics and prior knowledge. Since MOOC students have diverse background and motivation, there has been limited success in the self-selected or randomly assigned MOOC teams. Being part of an ineffective or dysfunctional team may well be inferior to independent study in promoting learning and can lead to frustration. This dissertation studies how to coordinate team based learning in MOOCs with a learning science concept, namely Transactivity. A transactive discussion is one where participants elaborate, build upon, question or argue against previously presented ideas [20]. It has long been established that transactive discussion is an important process that reflects good social dynamics in a group, correlates with students' increased learning, and results in collaborative knowledge integration. Building on this foundation, we design a deliberation-based team formation where students hold a course community deliberation before small group collaboration.

The center piece of this dissertation is a process for introducing online students into teams for effective group work. The key idea is that students should have the opportunity to interact meaningfully with the community before assignment into teams. That discussion not only provides evidence of which students would work well together, but it also provides students with a wealth of insight into alternative task-relevant perspectives to take with them into the collaboration.

The team formation process begins with individual work. The students post their individual work to a discussion forum for a community-wide deliberation over the work produced by each individual. The resulting data trace informs automated guidance for team formation. The automated team assignment process groups students who display successful team processes, i.e., where transactive reasoning has been exchanged during the deliberation. Our experimental results indicate that teams that are formed based on students' transactive discussion after the community deliberation have better collaboration product than randomly formed teams. Beyond team formation, this dissertation explores how to support teams in their teamwork after the teams have been formed. At this stage, in order to further increase a team's transactive communication during team work, we use an automated conversational agent to support team members' collaboration discussion through an extension of previously published facilitation techniques. As a grand finale to the dissertation, the paradigm for team formation validated in Amazon's Mechanical Turk is tested for external validity within two real MOOCs with different team-based learning setting. The results demonstrated the effectiveness of our team formation process.

This thesis provides a theoretical foundation, a hypothesis driven investigation both in the form of corpus studies and controlled experiments, and finally a demonstration of external validation. It's contribution to MOOC practitioners includes both practical design advice as well as coordinating tools for team based MOOCs.

August 30, 2016  
DRAFT

August 30, 2016  
DRAFT

# Contents

0.1	Terms . . . . .	1
<b>1</b>	<b>Introduction</b>	<b>3</b>
<b>2</b>	<b>Background</b>	<b>7</b>
2.1	Why Team-based Learning in MOOCs? . . . . .	7
2.1.1	Lack of social interactions in MOOCs . . . . .	7
2.1.2	Positive effects of social interaction on commitment . . . . .	7
2.1.3	Positive effects of social interaction on learning . . . . .	8
2.1.4	Desirability of team-based learning in MOOCs . . . . .	8
2.2	What Makes Successful Team-based Learning? . . . . .	9
2.3	Technology for Supporting Team-based Learning . . . . .	10
2.3.1	Supporting online team formation . . . . .	10
2.3.2	Supporting team process . . . . .	13
2.4	Discussion . . . . .	15
<b>3</b>	<b>Research Methodology</b>	<b>17</b>
3.1	Types of Studies . . . . .	17
3.2	Methods used in Corpus Analysis . . . . .	17
3.2.1	Text Classification . . . . .	17
3.2.2	Statistical Analysis . . . . .	19
3.3	Crowdsourced Studies . . . . .	20
3.3.1	Crowdsourced Experimental Design . . . . .	20
3.3.2	Collaborative Crowdsourcing . . . . .	21
3.3.3	Crowdworkers and MOOC students . . . . .	21
3.3.4	Constraint satisfaction algorithms . . . . .	22
3.4	Deployment Study . . . . .	22
3.4.1	How real MOOCs differ from MTurk . . . . .	22
<b>4</b>	<b>Factors that Correlated with Student Commitment in Massive Open Online Courses:</b>	
	<b>Study 1</b>	<b>25</b>
4.1	Introduction . . . . .	25
4.2	Coursera dataset . . . . .	27
4.3	Methods . . . . .	27
4.3.1	Learner Motivation . . . . .	27

4.3.2	Predicting Learner Motivation . . . . .	28
4.3.3	Level of Cognitive Engagement . . . . .	31
4.4	Validation Experiments . . . . .	33
4.4.1	Survival Model Design . . . . .	33
4.4.2	Survival Model Results . . . . .	34
4.4.3	Implications . . . . .	35
4.5	Chapter Discussions . . . . .	35
<b>5</b>	<b>Virtual Teams in Massive Open Online Courses: Study 2</b>	<b>39</b>
5.1	MOOC Datasets . . . . .	40
5.2	Study Groups in xMOOC Discussion Forums . . . . .	41
5.2.1	Effects of joining study groups in xMOOCs . . . . .	41
5.3	Virtual Team in NovoEd MOOCs . . . . .	43
5.3.1	The Nature of NovoEd Teams . . . . .	43
5.3.2	Effects of teammate dropout . . . . .	44
5.4	Predictive Factors of Team Performance . . . . .	45
5.4.1	Team Activity . . . . .	46
5.4.2	Communication Language . . . . .	46
5.4.3	Leadership Behaviors . . . . .	47
5.4.4	Results . . . . .	49
5.5	Chapter Discussions . . . . .	49
<b>6</b>	<b>Online Team Formation through a Deliberative Process in Crowdsourcing Environments: Study 3 and 4</b>	<b>51</b>
6.1	Introduction . . . . .	51
6.2	Method . . . . .	53
6.2.1	Experimental Paradigm . . . . .	53
6.3	Experiment 1. Group Transition Timing: Before Deliberation vs. After Deliberation . . . . .	58
6.3.1	Participants . . . . .	58
6.3.2	Results . . . . .	59
6.4	Experiment 2. Grouping Criteria: Random vs. Transactivity Maximization . . . . .	60
6.4.1	Participants . . . . .	61
6.4.2	Results . . . . .	61
6.5	Implications for team-based MOOCs . . . . .	63
6.6	Discussions . . . . .	64
6.7	Conclusions . . . . .	64
<b>7</b>	<b>Team Collaboration Communication Support: Study 5</b>	<b>67</b>
7.1	Bazaar Framework . . . . .	68
7.2	Prompts . . . . .	69
7.3	Learning Gain . . . . .	70
7.4	Participants . . . . .	70
7.5	Results . . . . .	71



7.5.1	Team Process and Team Performance . . . . .	71
7.5.2	Learning Gains . . . . .	72
7.6	Discussions . . . . .	72
7.6.1	High Transactivity Groups Benefit More from Bazaar Communication Support . . . . .	72
7.6.2	Learning Gain . . . . .	72
7.7	Conclusions . . . . .	72
<b>8</b>	<b>Team Formation Intervention Study in a MOOC: Study 6</b>	<b>73</b>
8.1	The Smithsonian Superhero MOOC . . . . .	73
8.2	Data Collection . . . . .	74
8.3	Adapting our Team Formation Paradigm to the MOOC . . . . .	75
8.3.1	Step 1. Individual Work . . . . .	75
8.3.2	Step 2. Course community deliberation . . . . .	75
8.3.3	Step 3. Team Formation and Collaboration . . . . .	75
8.4	Results . . . . .	76
8.4.1	Course completion . . . . .	76
8.4.2	Team collaboration communication . . . . .	78
8.4.3	Final project submissions . . . . .	78
8.4.4	Post-course Survey . . . . .	81
8.5	Discussion . . . . .	81
<b>9</b>	<b>Study 7. Team Collaboration as an Extension of the MOOC</b>	<b>85</b>
9.1	Research Question . . . . .	85
9.2	Adapting our team formation paradigm to this MOOC . . . . .	86
9.2.1	Individual Work . . . . .	86
9.2.2	Course Community Deliberation . . . . .	86
9.2.3	Team Formation . . . . .	86
9.2.4	Team Collaboration . . . . .	87
9.3	Results . . . . .	87
9.3.1	Team Collaboration and Communication . . . . .	87
9.3.2	Course Completion . . . . .	88
9.3.3	Team Project Submission . . . . .	89
9.3.4	Survey Results . . . . .	91
9.4	Discussions . . . . .	92
9.4.1	Problems with team space . . . . .	92
9.4.2	Team Member Dropout . . . . .	95
9.4.3	The Role of Social Learning in MOOCs . . . . .	95
9.5	Conclusion . . . . .	96
<b>10</b>	<b>Conclusions</b>	<b>97</b>
10.1	Reflections on the use of MTurk as a proxy for MOOCs . . . . .	97
10.2	Contributions . . . . .	98
10.3	Future Directions . . . . .	98

10.3.1	Compare with self selection based team formation . . . . .	98
10.3.2	Team Collaboration Support . . . . .	98
10.3.3	Community Deliberation Support . . . . .	99
10.3.4	Team-based Learning for Science MOOC . . . . .	99
<b>11</b>	<b>APPENDIX A: Final group outcome evaluation for Study 4</b>	<b>101</b>
<b>12</b>	<b>APPENDIX B. Survey for Team Track Students</b>	<b>105</b>
	<b>Bibliography</b>	<b>107</b>

# List of Figures

1.1	Thesis overview. . . . .	5
4.1	Study 1 Hypothesis. . . . .	26
4.2	Annotated motivation score distribution. . . . .	36
4.3	Survival curves for students with different levels of engagement in the Accountable Talk course. . . . .	37
4.4	Survival curves for students with different levels of engagement in the Fantasy and Science Fiction course. . . . .	37
4.5	Survival curves for students with different levels of engagement in the Learn to Program course. . . . .	38
5.1	Chapter hypothesis. . . . .	40
5.2	Screen shot of a study group subforum in a Coursera MOOC. . . . .	40
5.3	Homepage of a NovoEd team. 1: Team name, logo and description. 2: A team blog. 3: Blog comments. 4: Team membership roster. . . . .	43
5.4	Survival plots illustrating team influence in NovoEd. . . . .	45
5.5	Statistics of team size, average team score and average overall activity Gini coefficient. . . . .	46
5.6	Structural equation model with maximum likelihood estimates (standardized). All the significance level $p < 0.001$ . . . . .	47
6.1	Chapter hypothesis. . . . .	53
6.2	Collaboration task description. . . . .	54
6.3	Illustration of experimental procedure and worker synchronization for our experiment . . . . .	56
6.4	Workflow diagrams for Experiment 1. . . . .	58
6.5	Workflow diagrams for Experiment 2. . . . .	60
7.1	Chapter hypothesis. . . . .	68
7.2	Instructions for Pre and Post-test Task. . . . .	71
8.1	A screenshot of Week 4 discussion threads for team formation. . . . .	76
8.2	Instructions for Course Community Discussion and Team Formation. . . . .	77
8.3	Initial post in the team space. . . . .	78
8.4	Instructions for the final team project. . . . .	79

9.1	Initial Post. . . . .	87
9.2	Self introductions. . . . .	88
9.3	Voting for a team leader. . . . .	89
9.4	A brainstorming post. . . . .	90
9.5	Team communication frequency survey. . . . .	91
9.6	Team space list. . . . .	93
9.7	Team space in edX. . . . .	93
9.8	A to-do list. . . . .	94

# List of Tables

4.1	Features for predicting learner motivation. A binomial test is used to measure the feature distribution difference between the motivated and unmotivated post sets(**: $p < 0.01$ , ***: $p < 0.001$ ). . . . .	30
4.2	Accuracies of our three classifiers for the Accountable Talk course (Accountable) and the Fantasy and Science Fiction course (Fantasy), for in-domain and cross-domain settings. The random baseline performance is 50%. . . . .	31
4.3	Results of the survival analysis. . . . .	33
5.1	Statistics of the three Coursera MOOCs. . . . .	39
5.2	Statistics two NovoEd MOOCs. . . . .	40
5.3	Survival Analysis Results. . . . .	42
5.4	Survival analysis results. . . . .	44
5.5	Three different types of leadership behaviors. . . . .	48
6.1	Transactive vs. Non-transactive Discussions during Team Collaboration. . . . .	62
7.1	Number of teams in each experimental condition. . . . .	71
8.1	Student course completion across tracks. . . . .	76
8.2	Sample creative track projects. . . . .	80
8.3	Sample team track projects. . . . .	83
9.1	Student community deliberation in Week 1. . . . .	87
9.2	Finish status of the team projects. . . . .	89
11.1	Energy source X Requirement . . . . .	101
11.2	Energy X Energy Comparison . . . . .	102
11.3	Address additional valid requirements. . . . .	102
11.4	Address additional valid requirements. . . . .	103

August 30, 2016  
DRAFT

## 0.1 Terms

**Commitment:** in this thesis, commitment refers to whether participants stay in the MOOC instead of dropping out.

**Learning:** learning gain in this thesis is measured by the difference between pre and post-test.

**Team-based learning:** Team-based learning is one type of social learning approaches where students collaborate in a small team and accomplish a course project together. social learning approaches include Problem Based Learning, Team Project Based Learning, and Collaborative Reflection.

**Transactivity:** in this thesis, we focus on other-oriented transactivity, which is discussion that contains reasoning that operates on the reasoning of another [20].

**xMOOCs:** content-based MOOCs, such as Coursera and edX MOOCs.

August 30, 2016  
DRAFT



# Chapter 1

## Introduction

Ever since the emergent of MOOCs, many practitioners and platforms have been trying to implement team based learning in MOOCs. However, little success has been reported [206]. This thesis try to solve the team based learning challenges in MOOCs with a learning science concept of transactivity. A Transactive discussion is defined most simply as discussion that contains reasoning that “operates on the reasoning of another” [20]. This process is referred to alternately in the literature as argumentative knowledge construction [190], consensus building, transactive knowledge exchange, or even productive agency. The theoretical link between transactive reasoning and collaborative knowledge integration is based on Piaget’s proposal (1932/1965) that, when students operate on each other’s reasoning, they become aware of contradictions between their own reasoning and that of their partner. The neo-Piagetian perspective on learning shows that optimal learning between students occurs when students respect both their own ideas and those of the others that they are interacting with [51]. It has long been established that transactive discussion is an important process that reflects this kind of good social dynamics in a group. Transactive discussion is at the heart of what makes collaborative learning discussions valuable, leading to outcomes such as increased learning and collaborative knowledge integration [77].

Contemporary MOOCs are online learning environments where instructional design is based primarily on a purely instructivist approach. It features video lectures, quizzes, assignments, and exams as primary learning activities, with discussion forums as the only place that social interaction happens. The participation in the slow-motion forum communication is low in general [118]. Thus social isolation is still the norm of experience in the current generation of MOOCs. To improve students’ interaction and engagement, several MOOC platforms and recent research investigate how to introduce interactive peer learning into MOOCs. NovoEd (<https://novoed.com/>) is a smaller team-based MOOC platform that features a team-based, collaborative, and project-based approach. EdX is also releasing its new team based MOOC interface, which can be optionally incorporated into an instructivist MOOC. However, merely dropping the team based learning platform into a MOOC will not automatically lead to successful teams [103]. The percentage of active or successful teams is usually low across the team based MOOCs. There has been a lot of recent interest in designing short interventions with peer interaction in MOOCs, e.g. [41, 66, 103]. In these studies, MOOC participants are grouped together for a short period of time, e.g. 20 minutes, and discuss an assigned course related topic or a multiple choice question through synchronous text-based chat or Google Hangout. These studies have demonstrated

positive impact in terms of lower attrition immediately following the intervention or better performance in course assessments. Compared to these short-term, light weight social interaction interventions, team based learning offers an experience that is conducive for intense discussion to happen among students. Team based learning also provides students a unique opportunity to accomplish a course project together. Students can potentially gain more long term benefits compared to those short-term social interventions. It is more difficult to coordinate team based learning, especially in the online learning context. Without incorporation team based learning into the formal curriculum, most teams fail. This thesis designs and develops automated support for incorporating learning teams in a MOOC.

There are two main challenges associated with coordinating team based MOOCs: team formation and supporting formed teams. The two main approaches of forming online teams are self-selection and algorithm-based team formation. Self-selection is the current norm of team formation method adopted in recent team based MOOCs. However, MOOC students typically do not know each other. Very few students fill in their profile, making self-selection team formation difficult and inefficient in MOOCs. This thesis also provides evidence that many of these self-selected teams fail – there is very low activity level from the beginning. Using algorithm-based team formation, recent work applied clustering algorithm to MOOC team formation. Results showed that most of the teams had low activity level. Teams that were formed based on simple demographic information collected from surveys did not demonstrate more satisfactory collaboration compared to randomly formed teams [206]. The second challenge is how to coordinate team collaboration once the teams are formed. Students in the virtual learning teams usually have diverse backgrounds, a successful collaboration requires team members respect each other’s perspective and integrate each member’s unique knowledge. Owing to extremely skewed ratios of students to instructional staff, it can be prohibitively time-consuming for the instructional staff to manually coordinate all the formed teams. Therefore the second challenge of coordinating team based learning in MOOCs is how to automatically support teams that are already formed.

To leverage team-based learning, we need to address several problems that are associated with small team collaboration in MOOCs. Small teams often lose critical mass through attrition. The second danger of small teams is that students tend to depend too much on their team, lose intellectual connection with the community. Finally, since students rarely fill in their profiles or submit surveys, we have scant information of which students will work well together. To address these team formation challenges, we propose a deliberation-based team formation procedure, with transactive discussion as an integral part of the procedure, to improve the team initiation and selection process. Deliberation, or rational discourse that marshals evidence and arguments, can be effective for engaging groups of diverse individuals to achieve rich insight into complex problems [29]. In our team formation procedure, participants hold discussions in preparation for the collaboration task. We automatically assign teams in a way that maximizes average observed pairwise transactive discussion exchange within teams. By forming teams later, i.e. after community deliberation, we can avoid assigning students who are already dropping out into teams. By providing teams with the opportunity to join community deliberation before team collaboration, students may maintain the community connection. Moreover, we can leverage students’ transactive discussion evidence during the deliberation and group students who display successful team processes, i.e. transactive reasoning, in the deliberation. Based on the theoretical framing of the concept of transactivity and prior empirical work related to that construct, we

hypothesize that we can form more successful teams using students’ transactive discussion as evidence.

To support teams that are already formed, we study whether using collaboration communication support, i.e. conversational agent, can boost students’ transactive discussion during collaboration and enhance learning and collaboration product in online learning contexts. Previous research has demonstrated that conversational agents can positively impact team based learning through facilitating synchronous discussion [106]. In previous work on conversational agents, usually the agent facilitated discussion is a separate step before the collaboration step. We explore if we can combine these two steps, and support collaborative discussion while the students are working on the collaborative task.

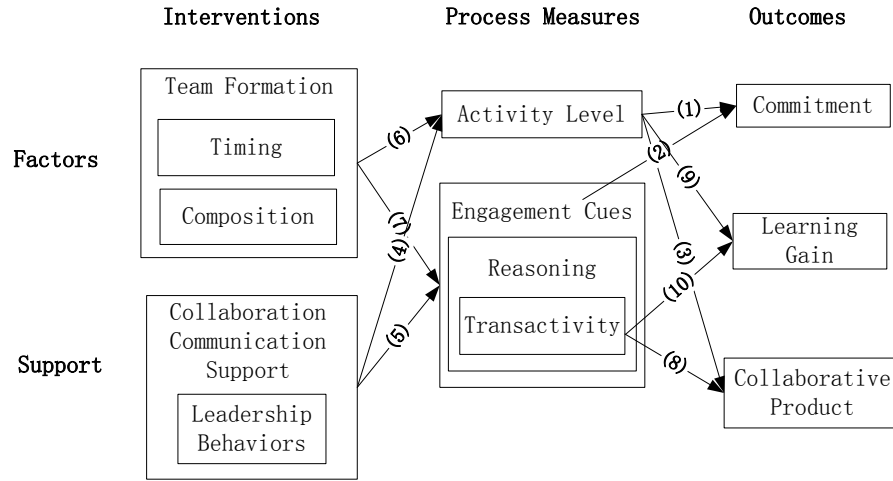


Figure 1.1: Thesis overview.

To sum up, Figure 1.1 presents a diagrammatic summary of our research hypotheses in this thesis. Given the importance of community deliberation in the MOOC context, we hypothesize that teams that formed after they engage in a large community forum deliberation process achieve better collaboration product than teams that instead perform an analogous deliberation within their small team. Based on the theories of transactive reasoning, we hypothesize that we can utilize transactive discussion as evidence to form teams that have more transactive reasoning during collaboration discussion ((7) in in Figure 1.1), and thus have better collaborative product in team based MOOCs ((8) in Figure 1.1). After teams are formed, we hypothesize that collaboration communication support can also lead to more transactive discussion during team collaboration ((5) in Figure 1.1), and thus lead to improved collaboration product ((8) in Figure 1.1).

This thesis provides an empirical foundation for proposing best practices in team based learning in MOOCs. In the following chapters I start with surveying related work on how to support online learning and improve collaboration product of online teams (Chapter 2). After the background chapter, I described research methodology of this thesis (Chapter 3). I begin with corpus analyses to form hypotheses (Chapter 4 and 5), then run controlled studies to confirm hypotheses (Chapter 6 and 7), and finally to apply the findings in two real MOOCs (Chapter 8 and 9).

This thesis presents four studies. Study 1 and 2 are corpus analyses with the goal of hypothesis formation. For any learning or collaboration to happen in the MOOC context, students need to stay engaged in the course. In **Study 1: Factors that are correlated with student commitment in xMOOCs (Chapter 4)**, we focused on predictive factors of commitment, which is the precondition to learning and collaboration. We found that students' activity level in the course forum were correlated with student commitment ((1) in Figure 1.1). We extracted linguistic features that indicate students' explicitly articulated reasoning from their discussion forum posts, and use survival models to validate that these features are associated with enhanced commitment in the course. ((2) in Figure 1.1).

In **Study 2: Behaviors that contribute to collaboration product in team based MOOCs (Chapter 5)**, we switch to study team based MOOCs, where a team based learning component is in the centre of a MOOC design. In particular, we study the collaborative interactions, i.e. asynchronous communication during virtual team collaboration, in team-based MOOC platform. Similar to Study 1, the activity level in a team, such as the number of messages sent between the team members each week, was positively correlated with team member commitment and collaboration product ((3) in Figure 1.1). We found that leadership behaviors in a MOOC virtual team were correlated with higher individual/group activity level and collaborative product quality ((4) and (5) in Figure 1.1). Results from Study 1 and 2 suggest that in order to improve collaborative product in team based MOOCs, it might be essential to focus on increasing activity level and the amount of reasoning discussion. We tackle this problem from two types of interventions: team formation and automatic collaboration discussion support for formed teams.

In **Study 3 and 4: Deliberation-based team formation support (Chapter 6)**, we ran controlled intervention studies in a crowdsourcing environment to confirm our team formation hypotheses. In Study 3, we confirmed the advantages of experiencing community wide transactive discussion prior to assignment into teams. In Study 4, we confirm that selection of teams in such a way as to maximize transactive reasoning based on observed interactions during the preparation task increases transactive reasoning during collaboration discussion and success in the collaborative product ((6) and (7) in Figure 1.1).

In **Study 5: Collaboration communication support (Chapter 7)**, we ran controlled intervention studies in the crowdsourcing environment using the same paradigm as Study 3 and 4. We study whether using collaboration communication support, i.e. conversational agent, can boost transactive reasoning during collaboration and enhance collaboration product and participants' learning gains ((8) in Figure 1.1).

The corpus analysis studies and crowdsourced experiments provide sufficient support for further investigation of team based learning in a field study. **Study 6 and 7: Deployment of the team formation support in real MOOCs (Chapter 8 and 9)** offers practical insights and experiences for how to incorporate a team based learning component in a typical xMOOC.

I have provided the ground work for supporting teams in MOOCs. This work has touched upon several topics of potential interest with future research directions. Chapter 10 provides a general discussion of the results, limitations of the work and a vision for future work.

# Chapter 2

## Background

In this chapter, we review background in three areas: positive effects of team-based learning in MOOCs, elements of successful team-based learning and efforts to support team-based learning.

### 2.1 Why Team-based Learning in MOOCs?

#### 2.1.1 Lack of social interactions in MOOCs

The current generation of MOOCs offer limited peer visibility and awareness. In most MOOCs, the only place that social interaction among student happen is course discussion forum. On the other hand, research on positive effects of collaboration on learning is a well-established field [210]. Many of the educationally valuable social interactions identified in this research are lost in MOOCs: online learners are “alone together” [178]. Analyses of attrition and learning in MOOCs both point to the importance of social engagement for motivational support and overcoming difficulties with material and course procedures. Recently, a major research effort has been started to explore the different aspects of social interaction amongst course participants [99]. The range of the examined options includes importing existing friend connections from social networks such as Facebook or Google+, adding social network features within the course platform, but also collaboration features such as discussion or learning groups and group exercises [159]. While this research has shown some initial benefits of incorporating more short term social interactions in MOOCs, less work has explored long term social interaction such as team-based learning.

#### 2.1.2 Positive effects of social interaction on commitment

The precondition of learning is students staying in the MOOC. Since MOOCs were introduced in 2011, there has been criticism of low student retention rates. While many factors contribute to attrition, high dropout rates are often attributed to feelings of isolation and lack of interactivity [40, 92] reasons directly relating to missing social engagement. One of the most consistent problems associated with distance learning environments is a sense of isolation due to lack of interaction [19, 80]. This sense of isolation is linked with attrition, instructional ineffectiveness,

failing academic achievement, and negative attitudes and overall dissatisfaction with the learning experience [23]. A strong sense of community has been identified as important for avoiding attrition [174]. In most current online classes, students opportunities for discussions with diverse peers are limited to threaded discussion forums. Such asynchronous text channels inhibit trust formation [140]. Attachment to sub-groups can build loyalty to the group or community as a whole [170]. Commitment is promoted when peer interaction is promoted [82].

More recent MOOCs try to incorporate scheduled and synchronous group discussion sessions where students are randomly assigned to groups [7, 66, 103]. Results show that having experienced a collaborative chat is associated with a slow down in the rate of attrition over time by a factor of two [176]. These studies indicate that small study groups or project teams in a bigger MOOC may help alleviate attrition.

### **2.1.3 Positive effects of social interaction on learning**

Peer learning is an educational practice in which students interact with other students to attain educational goals [51]. As Stahl and others note [158], Vygotsky argued that learning through interaction external to a learner precedes the internalization of knowledge. That interaction with the outside world can be situated within social interaction. So collaboration potentially provides the kind of interactive environment that precedes knowledge internalization. While peer learning in an online context has been studied in the field of CSCL, less is known about the effects of team based learning in MOOCs and how to support it.

There have been mixed effects reported by previous work on introducing social learning into MOOCs. Overall high satisfaction with co-located MOOC study groups that watch and study MOOC videos together has been reported [112]. Students who worked offline with someone during the course are reported to have learned 3 points more on average [26]. Coetzee et al. [42] tried encouraging unstructured discussion such as real-time chat room in a MOOC, but they did not find an improvement in students retention rate or academic achievement. Using the number of clicks on videos and the participation in discussion forums as control variables, Ferschke et al. [?] found that the participation in chats lowers the risk of dropout by approximately 50 %. More recent research suggest the most positive impact when experiencing a chat with exactly one partner rather than more or less in a MOOC [175]. The effect depends on how well the peer learning activities were incorporated into the curriculum. The effect depends also on whether sufficient support was offered to the groups.

### **2.1.4 Desirability of team-based learning in MOOCs**

There has been interest in incorporating a collaborative team-based learning component in MOOCs ever since the beginning. Benefits of group collaboration have been established by social constructivism [183] and connectivism [150]. Much prior work demonstrates the advantages of group learning over individual learning, both in terms of cognitive benefits as well as social benefits [15, 162]. The advantages of team-based Learning include improved attendance, increased pre-class preparation, better academic performance, and the development of interpersonal and team skills. Team-based learning is one of the few ways to achieve higher-level cognitive skills in large classes [122]. Social interactions amongst peers improves conceptual understanding

and engagement, in turn increasing course performance [48, 153] and completion rates [137]. NovoEd, a small MOOC platform, is designed specifically to support collaboration and project-based learning in MOOCs, through mechanisms such as participant reputational ranking, team formation, and non-anonymous peer reviews. The term GROOC has recently been defined, by Professor Mintzberg of McGill University (<https://www.mcgill.ca/desautels/programs/grooc>), to describe group-oriented MOOCs, based on one he has developed on social activism. edX is also releasing its new team based MOOC interface, which can be optionally incorporated into an instructivist MOOC. Limited success has been reported in these team-based MOOCs. Inspired by these prior work on team-based MOOC, in the next chapter we discuss what makes team-based learning successful, especially in online environment.

## 2.2 What Makes Successful Team-based Learning?

Although online students are “hungry for social interaction”, peer-based learning will not happen naturally without support [96]. In early MOOCs, discussion forums featured self-introductions from around the world, and students banded together for in-person meet-ups. Yet, when peer learning opportunities are provided, students don't always participate in pro-social ways; they may neglect to review their peers work, or fail to attend a discussion session that they signed up for; they may drop out from their team as they drop out from the course [96]. Learners have reported experiencing more frustrations in online groups than in face-to-face learning [152]. Many instructors assumed that a peer system would behave like an already-popular social networking service like Facebook where people come en masse at their own will. However, peer learning systems may need more active integration, otherwise, students can hardly benefit from them [159]. Providing the communication technological means is by far not sufficient. Social interaction cannot be taken for granted. E.g. in one MOOC that offers learning groups, only about 300 out of a total of 7,350 course participants joined one of the twelve learning groups [99]. The value of educational experiences is not immediately apparent to students, and those that are worthwhile need to be signaled as important in order to achieve adoption. Previous work found that offering even minimal course credit powerfully spurs initial participation [99].

The outcomes of small group collaborative activities are dependent upon the quality of collaborative processes that occur during the activity [17, 100, 157]. Specifically, lack of common ground between group members can hinder effective knowledge sharing and collective knowledge building [77], process losses [160] and a lack of perspective taking [101, 144, 145]. Transactivity is a property identified by many as an essential component to effective collaborative learning [51]. It is akin to discourse strategies identified within the organizational communication literature for solidarity building in work settings [139] as well as rhetorical strategies associated with showing openness [119]. The idea is part of the neo-Piagetian perspective on learning where it is understood that optimal learning between students occurs when students respect both their own ideas and those of the peers that they interact with, which is grounded in a balance of power within a social setting. Transactivity is known to be higher within groups where there is mutual respect [12] and a desire to build common ground [79]. High transactivity groups are associated with higher learning [90, 172], higher knowledge transfer [77], and better problem solving [12].

Automatic annotation of transactivity is not a new direction in the Computer Supported Collaborative Learning community. For example, several researchers have applied machine learning using text, such as newsgroup style interactions [143], chat data [90], and transcripts of whole group discussions [6]. Previous work in this research mostly study how to identify transactivity and its correlation with team process and success. This thesis, on the other hand, explores using identified transactivity as evidence for team formation.

## **2.3 Technology for Supporting Team-based Learning**

To support team-based learning in MOOCs, we can learn about effective support for social learning from the literature on Computer Supported Collaborative Learning. In a classroom, an instructor's role in promoting small group collaboration in the classroom includes 1) preparing students for collaborative work, 2) forming groups, 3) structuring the group-work task, and 4) influencing the student interaction process through teacher facilitation of group interaction [188]. In an online setting, some or all of this support may come through automated interventions. In this section, we survey the research on forming groups and facilitation of group interaction.

This chapter reviews technologies for supporting team formation and team process.

### **2.3.1 Supporting online team formation**

The existing research in team-based learning, building from a massive research base in traditional group work theory [43], has identified that group formation and maintenance require considerable extra planning and support. A group formation method is an important component for enhancing team member participation in small groups [85]. The three most common types of group forming methods are self selection, random assignment, and facilitator assignment [54].

#### **Self-selection based team formation**

Self selection, the most prevalent form of online grouping, is considered better for interaction but difficult to implement in a short time since in this case participants typically do not know each other and lack the face-to-face contact to “feel out” potential group members. For this reason, student teams in E-learning contexts are usually assigned by instructors, often randomly. Most current research supports instructor-formed teams over self-selected teams [132], although some authors disagree [14]. Self-selection has been recommended by some because it may offer higher initial cohesion [163], and cohesion has been linked to student team performance [71, 88, 198]. Self-selection in classrooms may have several pitfalls: a) Team formation around pre-existing friendships, which hampers the exchange of different ideas; b) The tendency of learners with similar abilities to flock together, so strong and weak learners do not mix, thus limiting interactions and preventing weaker learners to learn how stronger learners would tackle problems. The stronger learners would also not benefit from the possibilities to teach their peers, and c) Self-selection can also pose a problem for under-represented minorities. When an at-risk student is isolated in a group, this isolation can contribute to a larger sense of feeling alone. This can then lead to non-participation or purely passive roles [131].



## Algorithm-based team formation

Many team-formation problems is motivated by the following simply-stated, yet significant question: “Given a fixed population, how should one create a team (or teams) of individuals to achieve the optimal performance? [4]. To answer this question, an algorithm-based team formation needs two components: team formation criteria and team formation algorithm that optimizes that criteria.

### Team formation criteria

Most of the algorithm-based groups are formed based on information in learner’s profile. The fixed information may include: gender and culture. The information that require assessment include: expertise and learning styles [8, 50], knowledge about the content, personality, attributes [72], the current (and previous) learners goals [133], knowledge and skills [74], the roles and strategies used by learners to interact among themselves and teachers preferences (See [133] for an overview). Because mixed-gender groups consisting of students with different cognitive styles would establish a community with different ideas, so that students could see and design their products from different angles and provide solutions from different perspectives, which might help them to achieve better performance during CSCL. However, the results varied when it comes to whether to form a heterogeneous or homogeneous group. Several attempts have been made to analyze the effects of gender grouping on students group performance in CSCL, but the findings to date have been varied [204]. Founded on the Piagetian theory of equilibrium, Doise and Mugny understand the socio-cognitive conflict as a social interaction induced by a confrontation of divergent solutions from the participant subjects. From that interaction the individuals can reach higher states of knowledge[57]. Paredes et al. [134] shows that heterogeneous groups regarding learning styles tend to perform better than groups formed by students with similar characteristics. For example, by combining students according to two learning style dimensions: active/reflective and sequential/global, the heterogeneous groups have more interaction during collaboration [55]. While homogeneous groups are better at achieving specific aims, heterogeneous groups are better in a broader range of tasks when the members are balanced in terms of diversity based on functional roles or personality differences [130]. It’s not the heterogeneous vs. homogeneous distinction that leads to the different results, it’s the interaction happens among teams which can also be influenced by other factors like whether instructors have designed activities that suit the teams, language, culture, interests, individual personalities [75]. Recent work shows that balancing for personality leads to significantly better performance on a collaborative task [116]. Since students in MOOC have much more diverse background and motivation compared with traditional classroom, we think a more useful criteria for team formation is evidence of how well they are already interacting or working with each other in the course. Besides leveraging the static information from student profiles, recent research tries to include dynamic inputs from the environment to form teams, such as the availability of specific tools and learning materials [196], or emotional parameters of learners [211], learner context information such as location, time, and availability [65, 126]. Isotani et al. [87] proposed a team formation strategy to first understand students needs (individual goals) and then select a theory (and also group goals) to form a group and design activities that satisfy the needs of all students within a group.

It is less studied how to form teams without learner profile or previously existed links between

participants. In MOOCs, there is very limited student profile information. Recently, Zheng et al. [206] shows that teams that were formed based on simple demographic information collected from surveys did not demonstrate more satisfactory collaboration compared to randomly formed teams. To form effective teams in environment where participants have no history of communication or collaboration before, Lykourantzou et al. explores a strategy for “team dating”, a self-organized crowd team formation approach where workers try out and rate different candidate partners. They find that team dating affects the way that people select partners and how they evaluate them [117]. The proposed team formation strategy in this thesis tries to group students based on dynamic evidence extracted from students’ interaction.

### **Team formation algorithms**

Clustering algorithms constitute the most applied techniques for automatic group formation (e.g. [10, 189]). Methods such as C-Means, K-means and EM have all demonstrated success [38]. The clustering is usually done based on students knowledge or errors observed in an individual task (e.g. [189]). Heterogeneous and homogeneous groups can be formed after clustering. Most of the algorithms were evaluated

Team formation can also be formulated as a constraint satisfaction problem. Given a task  $T$ , a pool of individuals  $X$  with different skills, and a social network  $G$  that captures the compatibility among these individuals, Lappas et al. [108] studies the problem of finding a subset of  $X$ , to perform the task. The requirement is that members of  $X$  not only meet the skill requirements of the task, but can also work effectively together as a team. The existence of an edge (or the shortest path) between two nodes in  $G$  indicates that the corresponding persons can collaborate effectively. Anagnostopoulos et al. [9] proposes an algorithm to form online teams for a certain team given the social network information. (i) each team possesses all skills required by the task, (ii) each team has small communication overhead, and (iii) the workload of performing the tasks is balanced among people in the fairest possible way. Group fitness is a measure of the quality of a group with respect to the group formation criteria, and partition fitness is a measure of the quality of the entire partition. Many team formation algorithms are evaluated on generated or simulated data from different distributions [4].

Ikeda et al. [84] proposed an “opportunistic group formation”. Opportunistic Group Formation is the function to form a collaborative learning group dynamically. When it detects the situation for a learner to shift from individual learning mode to collaborative learning mode, it forms a learning group each of whose members is assigned a reasonable learning goal and a social role which are consistent with the goal for the whole group. Unfortunately, there is no literature on the architecture of the developed systems or their evaluation. Following this work, opportunistic collaboration, in which groups form, break up, and recombine as part of an emerging process, with all participants aware of and helping to advance the structure of the whole [205]. Once the opportunistic group formation model finds that the situation of a learner is the right timing to shift the learning mode from individual learning to the collaborative learning, the system taking charge of the learner proposes other systems to begin the negotiation for form the learning group formation. The opportunistic group formation system, called I-MINDS, was evaluated against the performance of the teams, measured based on the teams outcomes and their responses to a series of questionnaires that evaluates team-based efficacy, peer rating, and individual evaluation [155]. The results showed that students using I-MINDS performed (and outperformed in some aspects) as well as students in face-to-face settings.

## Group formation in social networks

Because students in MOOCs typically do not have existing ties prior to group formation, team formation in MOOCs is similar to group formation in social networks, especially online creative communities where participants form teams to accomplish projects. The processes by which communities and online groups come together, attract new members, and develop over time is a central research issue in the social sciences [45]. Previous work study how the evolution of these communities relates to properties such as the structure of the underlying social networks (e.g. friendship link [13], co-authorship link [148]). The findings include the tendency of an individual to join a community or group is influenced not just by the number of friends he or she has within the community, but also crucially by how those friends are connected to one another. In addition to communication, shared interests and status within the group are key predictors of whether two individuals will decide to work together. Theories on social exchange help lay a foundation for how communication affect people’s desire to collaborate as well as their ultimate success [61]. Research suggests that communication plays a key role in how online relationships form and function. In a study of online newsgroups, McKenna et al. [121] found communication frequency to be a significant factor in how relationships develop, and in whether these relationships persist years later. Communication not only helps relationships to form, but it improves working relationships in existing teams as well. For example, rich communication helps to support idea generation [166], creation of shared mental models [68], and critique exchange [146] in design teams. Based on this research, it is reasonable to believe that participants who showed substantial discussion with each other may develop better collaborative relationship in small team.

Given a complex task requiring a specific set of skills, it is useful to form a team of experts who work in a collaborative manner against time and many different costs. For example, we need to find a suitable team to answer community-based questions and collaboratively develop software. Or to find a team of well-known experts to review a paper. This team formation problem needs to consider factors like communication overhead and load balancing. Wang et al. [185] surveyed the state-of-the-art team formation algorithms. The algorithms were evaluated on datasets such as DBLP, IMDB, Bibsonomy and StackOverflow based on different metrics such as communication cost. Unfortunately, these algorithms were rarely evaluated in real team formation and team collaboration tasks.

### 2.3.2 Supporting team process

Team process support is important for the well functioning of the team once the it is formed. A conceptual framework for team process support referred to as the Collaboration Management Cycle is studied in [156]. This foundational work was influential in forming a vision for work on dynamic support for collaborative learning. In this work, Soller and colleagues provided an ontology for types of support for collaborative learning. They illustrated the central role of the assessment of group processes underlying the support approaches, including (a) mirroring tools that reflect the state of the collaboration directly to groups, (b) meta-cognitive tools that engage groups in the process of comparing the state of their collaboration to an idealized state in order to trigger reflection and planning for improvement of group processes, and finally, (c)

guiding systems that offer advice and guidance to groups. At the time, guiding systems were in their infancy and all of the systems reviewed were research prototypes, mostly not evaluated in realistic learning environments.

### **Collaboration support systems**

There are in general three types of collaboration support systems. Mirroring systems, which display basic actions to collaborators. For example, chat awareness tools such as Chat Circles [180], can help users keep track of ongoing conversations. Metacognitive tools, which represent the state of interaction via a set of key indicators. Talavera and Gaudioso [168] apply data mining and machine learning methods to analyze student messages in asynchronous forum discussions. Their aim is to identify the variables that characterize the behaviour of students and groups by discovering clusters of students with similar behaviours. A teacher might use this kind of information to develop student profiles and form groups. Anaya et al. [10] classify and cluster students based on their collaboration level during collaboration and show this information to both teachers and students. Coaching systems, which offer advice based on an interpretation of those indicators. OXEnTCHE [181] is an example of a sentence opener-based tool integrated with an automatic dialogue classifier that analyses on-line interaction and provides just-in-time feedback (e.g. productive or non-productive) to both teachers and learners. Fundamentally, these three approaches rely on the same model of interaction regulation, in that first data is collected, then indicators are computed to build a model of interaction that represents the current state, and finally, some decisions are made about how to proceed based on a comparison of the current state with some desired state. Using social visualizations to improve group dynamics is a powerful approach to supporting teamwork [18]. Another technique has been to use specifically crafted instructions to change group dynamics [120]. Previous work monitors the communication pattern in discussion groups with real-time language feedback [169]. The difference between the three approaches above lies in the locus of processing. Systems that collect interaction data and construct visualizations of this data place the locus of processing at the user level, whereas systems that offer advice process this information, taking over the diagnosis of the situation and offering guidance as the output. In the latter case, the locus of processing is entirely on the system side. Less research studied coordinating longitudinal virtual teams, we investigate how students collaborate through a longer duration of an entire MOOC in NovoEd, where small group collaborations are required. We also explore how to support teams where team collaboration is optional in xMOOCs.

### **Collaboration discussion process support**

The field of Computer Supported Collaborative Learning (CSCL) has a rich history extending for over two decades, covering a broad spectrum of research related to facilitation collaborative discussion and learning in groups, especially in computer mediated environments. A detailed history is beyond the scope of this thesis, but interested readers can refer to Stahls well known history of the field [157]. An important goal of this line of research is to develop environments with affordances that support successful collaboration discussion. And a successful collaborative discussion is where team members show respect to each other's perspective and operate on each

other’s reasoning.

Gweon et al. [77] identified five different group processes that instructors believe are important in accomplishing successful group work. The processes are goal setting, progress, knowledge co-construction, participation, and teamwork. They also automatically monitors group processes using natural language processing and machine learning. Walker [184] and Kumar & Ros [104] were the first to develop full-fledged guiding systems that have been evaluated in large scale studies in real classrooms. While there are many technical differences between the Walker [184] and Kumar et al. [104] architectures, what they have in common is the application of machine learning to the assessment of group processes. This thesis also utilize machine learning models to automatically identify transactive discussions from students’ deliberation.

Alternative perspectives during collaboration discussion can stimulate students’ reflection [51]. Therefore students can benefit to some extent from working with another student, even in the absence of scaffolding [78, 106]. Research in Computer-Supported Collaborative Learning has demonstrated that conversational computer agents can serve as effective automated facilitators of synchronous collaborative learning [59]. Conversational agent technology is a paradigm for creating a social environment in online groups that is conducive to effective teamwork. Kumar and Rose [105] has demonstrated advantages in terms of learning gains and satisfaction scores when groups learning together online have been supported by conversational agents that employ Balesian social strategies. Previous work also shows that context sensitive support for collaboration is more effective than static support [106]. Personalized agents further increase supportiveness and help exchange between students [106]. Students are sensitive to agent rhetorical strategies such as displayed bias, displayed openness to alternative perspectives [104], and targeted elicitation [83]. Accountable talk agents were designed to intensify transactive knowledge construction, in support of group knowledge integration and consensus building [190, 191]. Adamson [1] investigates the use of adaptable conversational agents to scaffold online collaborative learning discussions through an approach called Academically Productive Talk (APT), or Accountable Talk. Their results demonstrates that APT based support for collaborative learning can significantly increase learning, but that the effect of specific APT facilitation strategies is context-specific. Rosé and Ferschke studies supporting discussion-based learning at scale in MOOCs with Bazaar Collaborative Reflection, making synchronous collaboration opportunities available to students in a MOOC context. Using the number of clicks on videos and the participation in discussion forums as control variables, they found that the participation in chats lowers the risk of dropout by approximately 50% [66].

## 2.4 Discussion

There are many potential ways to enhance MOOC students’ learning and positively influence student interactions, the first section of this chapter shows that it is desirable to incorporate team-based learning in the MOOC context. Students in MOOCs typically do not have existing ties prior to group formation. Most previous team formation work will not work in the context of MOOCs. Based on previous team formation research, we propose a forum deliberation process into the team formation process. We hypothesize that the communication in the forum deliberation can be utilized as evidence for group formation. Informed by the research on transactivity,

we study whether teams composed of individuals with a history of engaging in more transactive communication during a pre-collaboration deliberation achieve more effective collaboration in their teams. MOOC students often have various backgrounds and perspectives, we leverage a conversational agent to encourage students' overt reasoning of conflict perspectives and further boost the transactive discussion during team discussion. Building on the paradigm for dynamic support for team process that has proven effective for improving interaction and learning in a series of online group learning studies, our conversational agent uses some Accountable talk moves to encourage students to reason from different perspectives during discussion as it unfolds [2, 106, 107].

## Chapter 3

# Research Methodology

This thesis utilized three types of research methods: corpus data analysis, crowdsourced study and field deployment study.

### 3.1 Types of Studies

In order for the research to practically solve real problems in MOOCs, we perform different types of studies in sequence. We first perform corpus analysis studies on data collected in previous MOOCs to form hypotheses. Corpus analysis is good for establishing correlational relationships between variables. To test our hypotheses, we run controlled studies in a crowdsourced environment. In crowdsourced studies, mechanical workers take the role of students. The experiments can be similar to lab studies. But since it is much easier to recruit workers in crowdsourced environments, it enables us to quickly iterate on the experiments design and draw causal relationships between variables. Finally, to understand other contextual factors that may play an important role, we apply the findings in the deployment study in a real MOOC.

### 3.2 Methods used in Corpus Analysis

The corpus analysis studies analyze and model different forms of social interactions across several online environments. A commonly used method is text classification methods. We describe how we apply text classification to social interaction analysis in Study 1 and 2, where we adopt a content analysis approach to analyze students social interactions in MOOCs.

#### 3.2.1 Text Classification

Text classification is an application of machine learning technology to a structured representation of text. The studies in this thesis utilize text classification to automatically annotate each message or conversational turn with a certain measure, such as transactivity. Machine learning algorithms can learn mappings between a set of input features and a set of output categories. They do this by examining a set of hand coded “training examples” that exemplify each of the target categories.

The goal of the algorithm is to learn rules by generalizing from these examples in such a way that the rules can be applied effectively to new examples.

## **Social Interaction Analysis in MOOCs**

The process-oriented research on collaborative learning is faced with an enormous amount of data. Applications of machine learning to automatic collaborative-learning process analysis are growing in popularity within the computer-supported collaborative learning (CSCL) community [79]. Previous work has analyzed content of student dialogues in tutoring and computer-supported collaborative learning environments. Chi [31] pointed out the importance of verbal analysis, which is a way to indirectly view student cognitive activity. De Wever [52] further demonstrated that content analysis has the potential to reveal deep insights about psychological processes that are not situated at the surface of internalized collaboration scripts. Previous work in the field of CSCL (Computer Supported Collaborative Learning) has demonstrated that discussion can facilitate learning in traditional contexts such as classrooms, and intelligent tutoring systems [33]. In most current xMOOCs, the only social interactions between students are threaded discussions in course forums. Unlike traditional education settings, discussions in xMOOCs are large-scaled and asynchronous in nature, and thereby more difficult to control. Many student behaviors have been observed in discussion forums, e.g., question answering, self-introduction, complaining about difficulties and corresponding exchange of social support. A very coarse grained distinction in posts could be on vs. off topic.

In Study 1, significant connection that has been discovered between linguistic markers extracted from discussion posts in MOOC forums and student commitment. In Study 2, we identified the leadership behaviors from the communications among team members in NovoEd MOOCs. These behaviors were found to be correlated with team final task performance. In Study 3 and 4, we reveal the potential of utilizing transactivity analysis for forming effective crowd worker or student teams. Since transactivity is a sign of common ground and team cohesiveness building, in Study 3, we predict transactivity based on hand-code data based on a well-established framework from earlier research [8] in an attempt to capture crowd workers discussion behaviors. We used the simplest text classification method to predict transactivity since the variance of crowd workers' discussion posts was low. Few of the posts were off-topic. And because we specifically asked the Turkers to give feedback transactively, more than 70% of the posts turned out to be transactive, which is much higher than a typical forum discussion. For a different domain or context, more complicated model and features will be necessary to achieve a reasonable performance.

## **Collaborative Learning Process Analysis**

It has been long acknowledged that conversation is a significant way for students to construct knowledge and learn. Previous studies on learning and tutoring systems have provided evidence that students participation in discussion is correlated with their learning gains [16, 35, 44]. Social interactions that are meaningful to learning from two aspects:

(1) The cognitive aspects such as the reasoning that is involved in the discussion. Students cognitively relevant behaviors, which are associated with important cognitive processes that pre-



cede learning may be found in discussion forums. Research has consistently found that the cognitive processes involved in higher-order thinking lead to better knowledge acquisition [32, 34, 73]. Previous work has investigated students' cognitive engagement in both face-to-face [46] and computer mediated communication (CMC) environments [69, 208]. In this chapter, we try to measure the cognitive engagement of a MOOC user based on how much personal interpretation are contained in the posts.

(2) The social aspects. During social interactions, students pay attention to each other which is important for building common ground and team cohesiveness later in their team collaboration. The social modes of transactivity describe to what extent learners refer to contributions of their learning partners, which has been found to be related to knowledge acquisition [67, 171].

In order to tie these two aspects together, we use the measure of transactivity to evaluate the value associated with social interactions. There are a variety of subtly different definitions of transactivity in the literature, however, they frequently share the two aspects: the cognitive aspect which requires reasoning to be explicitly displayed in some form, and the social aspect which requires connections to be made between the perspective of one student and that of another.

The area of automatic collaborative process analysis has focused on discussion processes associated with knowledge integration, where interaction processes are examined in detail. The analysis is typically undertaken by assigning coding categories and counting pre-defined features. Some of the research focus on counting the frequency of specific speech acts. However, speech acts may not well represent relevant cognitive processes of learning. Frameworks for analysis of group knowledge building are plentiful and include examples such as Transactivity [20, 171, 190] Intersubjective Meaning Making [165], and Productive Agency [147]. Despite differences in orientation between the cognitive and socio-cultural learning communities, the conversational behaviors that have been identified as valuable are very similar. Schwartz and colleagues [147] and de Lisi and Golbeck [51] make very similar arguments for the significance of these behaviors from the Vygotskian and Piagetian theoretical frameworks respectively.

In this thesis we are focusing specifically on transactivity. More specifically, our operationalization of transactivity is defined as the process of building on an idea expressed earlier in a conversation using a reasoning statement. Research has shown that such knowledge integration processes provide opportunities for cognitive conflict to be triggered within group interactions, which may eventually result in cognitive restructuring and learning [51]. While the value of this general class of processes in the learning sciences has largely been argued from a cognitive perspective, these processes undoubtedly have a social component, which we explain below and use to motivate our technical approach.

### **3.2.2 Statistical Analysis**

#### **Survival Models**

Survival models can be regarded as a type of regression model, which captures influences on time-related outcomes, such as whether or when an event occurs. In our case, we are investigating our engagement measures' influence on when a course participant drops out of the course forum. More specifically, our goal is to understand whether our automatic measures of student engagement can predict her length of participation in the course forum. Survival analysis

is known to provide less biased estimates than simpler techniques (e.g., standard least squares linear regression) that do not take into account the potentially truncated nature of time-to-event data (e.g., users who had not yet left the community at the time of the analysis but might at some point subsequently). From a more technical perspective, a survival model is a form of proportional odds logistic regression, where a prediction about the likelihood of a failure occurring is made at each time point based on the presence of some set of predictors. The estimated weights on the predictors are referred to as hazard ratios. The hazard ratio of a predictor indicates how the relative likelihood of the failure occurring increases or decreases with an increase or decrease in the associated predictor. We use the statistical software package Stata [47]. We assume a Weibull distribution of survival times, which is generally appropriate for modeling survival. Effects are reported in terms of the hazard ratio, which is the effect of an explanatory variable on the risk or probability of participants drop out from the course forum. Because the Activity variable has been standardized, the hazard rate here is the predicted change in the probability of dropout from the course forum for a unit increase in the predictor variable (i.e., binary variable changing from 0 to 1 or the continuous variable increasing by a standard deviation when all the other variables are at their mean levels).

In Study 1, we use survival models to understand how attrition happens along the way as students participate in a course. This approach has been applied to online medical support communities to quantify the impact of receipt of emotional and informational support on user commitment to the community [187]. Yang et al. [200] and Rose et al. [142] have also used survival models to measure the influence of social positioning factors on drop out of a MOOC. Study 1 belongs to this body of work.

## Structural Equation Models

A Structural Equation Model [22], is a statistical technique for testing and estimating correlational (and sometimes causal) relations in cross sectional datasets. In Study 2, to explore the influence of latent factors, we take advantage of SEM to formalize the conceptual structure in order to measure what contributes to team collaboration product quality.

## 3.3 Crowdsourced Studies

### 3.3.1 Crowdsourced Experimental Design

Instructors may be reluctant to deploy new instructional tools that students dislike. Rather than trying out untested designs on real live courses, we have prototyped and tested the approach using a crowdsourcing service, Amazon Mechanical Turk, (MTurk). Crowdsourcing is a way to quickly access a large user pool, collect data at a low cost [94]. The power and the generality of the findings obtained through empirical studies are bounded by the number and type of participating subjects. In MTurk, it is easier to do large scale studies. Researchers in other fields have used crowdsourcing to evaluate designs [81].

MTurk provides a convenient labor pool and deployment mechanism for conducting formal experiments. For a factorial design, each cell of the experiment can be published as an individual

HIT and the number of responses per cell can be controlled by throttling the number of assignments. Qualification tasks may optionally be used to enforce practice trials and careful reading of experimental procedures. The standard MTurk interface provides a markup language supporting the presentation of text, images, movies, and form-based responses; however, experimenters can include interactive stimuli by serving up their own web pages that are then presented on the MTurk site within an embedded frame.

As with any experimental setting, operating within an infrastructure like Mechanical Turk poses some constraints that must be clearly understood and mitigated. For example, one of such constraints is for the study to be broken into a series of tasks. Even then, our study uses tasks that are much more complex and time consuming (our task takes about an hour) than those recommended by the Mechanical Turk guidelines. Other constraints are the ability to capture the context under which the participants completed the tasks, which can be a powerful factor in the results, and the ability to collect certain metrics for the tasks being performed by participants, most specifically the time to completion. While time to completion is reported by Mechanical Turk in the result sets, it may not accurately measure the task time in part because we do not get a sense for how much time the user spent on the task and how much time the task was simply open in the browser. We note, however, that many of the design limitations can be mitigated by using Mechanical Turk as a front end to recruit users and manage payment, while implementing the actual study at a third-party site and including a pointer to that site within a Mechanical Turk HIT. Once the user finishes the activity at the site, they could collect a token and provide it to Mechanical Turk to complete a HIT. Clearly, this involves additional effort since some of the support services of the infrastructure are not used, but the access to the large pool of users to crowdsource the study still remains.

### **3.3.2 Collaborative Crowdsourcing**

Collaborative crowdsourcing, i.e. the type of crowdsourcing that relies on teamwork, is often used for tasks like product design, idea brainstorming or knowledge development. Prior systems have shown that multiple workers can be recruited for collaboration by having workers wait until a sufficient number of workers have arrived [36, 116, 182]. Although an effective team formation method may enable crowds to do more complex and creative work, forming teams in a way that optimizes outcomes is a comparably new area for research [117].

### **3.3.3 Crowdworkers and MOOC students**

Mechanical workers have different motivation compared with MOOC students. Some concerns, such as subject motivation and expertise, apply to any study and have been previously investigated. Heer and Bostock [81] has replicated prior laboratory studies on spatial data encodings in MTurk. Kittur et al. [94] used MTurk for collecting quality judgments of Wikipedia articles. Turker ratings correlated with those of Wikipedia administrators. As a step towards effective team-based learning in MOOCs, in this Study 3 and 4, we explore the team-formation process and team collaboration support in an experimental study conducted in an online setting that enables effective isolation of variables, namely Amazon's Mechanical Turk (MTurk). While crowd workers likely have different motivations from MOOC students, their remote individual work

setting without peer contact resembles today’s MOOC setting where most students learn in isolation [41]. The crowdsourced environment enables us to quickly iterate on our experimental design. This allows us to test the causal connection between variables in order to identify principles that later we will test in an actual MOOC. A similar approach was taken in prior work to inform design of MOOC interventions for online group learning [41]. Rather than trying out untested designs on real live courses, we think it is necessary to prototype and test the approach first using the crowdsourcing environment. When designing deployment studies in real MOOCs, there will likely be constraints and other factors that are hard to control. By doing controlled crowdsourced study, we will be able to directly answer our research questions.

### 3.3.4 Constraint satisfaction algorithms

#### Minimal Cost Max Network Flow algorithm

For the Transactivity Maximization condition, teams were formed so that the amount of transactive discussion among the team members was maximized. A Minimal Cost Max Network Flow algorithm was used to perform this constraint satisfaction process [5]. This standard network flow algorithm tackles resource allocation problems with constraints. In our case, the constraint was that each group should contain four people who have read about different energies (i.e. a Jigsaw group). At the same time, the minimal cost part of the algorithm maximized the transactive communication that was observed among the group members during Deliberation step. The algorithm finds an approximately optimal grouping within  $O(N^3)$  ( $N$  = number of workers) time complexity. A brute force search algorithm, which has an  $O(N!)$  time complexity, would take too long to finish since the algorithm needs to operate in real time. Except for the grouping algorithm, all the steps and instructions were identical for the two conditions.

## 3.4 Deployment Study

To evaluate our designs and study how applicable the results may be to MOOCs, we need to do deployment study. The field study method inherently lacks the control and precision that could be attained in a controlled setting [202, 203]. There is also a sampling bias given that all of our participants self-selected into our team track. The way the interview data was collected and coded affects its interpretation and there may be alternative explanations.

### 3.4.1 How real MOOCs differ from MTurk

There are three key difference between MOOC and MTurk. First, crowd workers likely have different motivations from MOOC students. The dropout rate of MOOC students is much higher than crowdsourced workers. It’s important to study the intervention’s impact on students’ dropout in a real MOOC deployment study. It is less interesting to study the effect on dropout in MTurk. Second, instructors may be reluctant to deploy new instructional tools that students dislike. Only a deployment study can answer questions like: how many students are interested in adopting the intervention design? Whether the intervention design is enjoyable or not? Third,

August 30, 2016

DRAFT

an MTurk task usually lasts less than an hour. For our team formation intervention, the team collaboration can be several weeks. A deployment study will help us understand what support is needed during this longer period of time.

Before the actual deployment study in the MOOC, we customize our design around the need of the MOOC instructors and students. The designs and methods need to be adapted to the content of the course.



## Chapter 4

# Factors that Correlated with Student Commitment in Massive Open Online Courses: Study 1

1

For any learning or collaboration to happen in the MOOC context, students need to stay engaged in the course. In this chapter, we explored what factors are correlated with student commitment, which is the precondition for learning and collaboration product. Students' activity level in the course, e.g. the number of videos a student watches each week, is correlated with student's commitment. Learner engagement cues automatically extracted from the text of forum posts, including motivational cues and overt reasoning behaviors, are also correlated with student's commitment. We validate these factors using survival models that evaluate the predictive validity of these variables in connection with attrition over time. We conduct this evaluation in three MOOCs focusing on very different types of learning materials.

Previous studies on learning and tutoring systems have provided evidence that students participation in discussion [e.g., 2, 9, 12] is correlated with their learning gains in other instructional contexts. Brinton et al. [27] demonstrates that participation in the discussion forums at all is a strong indicator of student commitment. Wang et al. showed that students cognitively relevant behaviors in a MOOC discussion forum is correlated with their learning gains. [186]. We hypothesize that engagement cues extracted from discussion behaviors in MOOC forums are correlated with student commitment and learning. Figure 4.1 presents our hypotheses in this chapter.

### 4.1 Introduction

High attrition rate has been a major criticism of MOOCs. Only 5% students who enroll in MOOCs actually finish [98]. In order to understand students' commitment, especially given the varied backgrounds and motivations of students who choose to enroll in a MOOC, [53], we gauge a student's engagement using linguistic analysis applied to the student's forum posts

<sup>1</sup>The contents of this chapter are modified from three published papers: [193], [194] and [192]

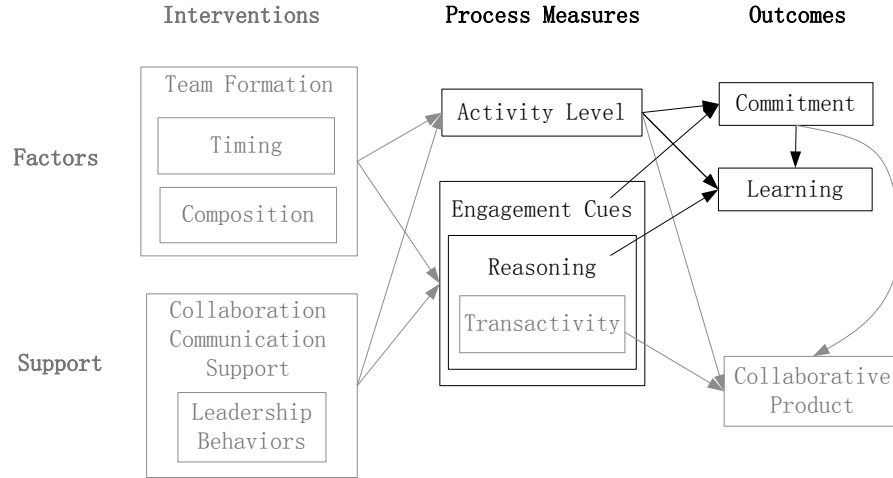


Figure 4.1: Study 1 Hypothesis.

within the MOOC course. Based on the learning sciences literature, we quantify a student’s level of engagement in a MOOC from two different angles: (1) displayed level of motivation to continue with the course and (2) the level of cognitive engagement with the learning material. Student motivation to continue is important- without it, it is impossible for a student to regulate him or her effort to move forward productively in the course. Nevertheless, for learning it is necessary for the student to process the course content in a meaningful way. In other words, cognitive engagement is critical. Ultimately it is this grappling with the course content over time that will be the vehicle through which the student achieves the desired learning outcomes.

Conversation in the course forum is replete with terms that imply learner motivation. These terms may include those suggested by the literature on learner motivation or simply from our everyday language. E.g. “I tried very hard to follow the course schedule” and “I couldn’t even finish the second lecture.” In this chapter, we attempt to automatically measure learner motivation based on such markers found in posts on the course discussion forums. Our analysis offers new insights into the relation between language use and underlying learner motivation in a MOOC context.

Besides student motivational state, the level of cognitive engagement is another important aspect of student participation [28]. For example, “This week’s video lecture is interesting, the boy in the middle seemed tired, yawning and so on.” and “The video shows a classroom culture where the kids clearly understand the rules of conversation and acknowledge each others contribution.” These two posts comment on the same video lecture, but the first post is more descriptive at a surface level while the second one is more interpretive, and displays more reflection. We measure this difference in cognitive engagement with an estimated level of language abstraction. We find that users whose posts show a higher level of cognitive engagement are more likely to continue participating in the forum discussion.

The distant nature and the sheer size of MOOCs require new approaches for providing student feedback and guiding instructor intervention [138]. One big challenge is that MOOCs are far from uniform. In this chapter, we test the generality of our measures in three Coursera MOOCs focusing on distinct subjects. We demonstrate that our measures of engagement are consistently



predicative of student dropout from the course forum across the three courses. With this validation, our hope is that in the long run, our automatic engagement measures can help instructors target their attention to those who show serious intention of finishing the course, but nevertheless struggle through due to dips in learner motivation. Our linguistic analysis provides further indicators that some students are going through the motions in a course but may need support in order to fully engage with the material. Again such monitoring might aid instructors in using their limited human resources to the best advantage.

## 4.2 Coursera dataset

In preparation for a partnership with an instructor team for a Coursera MOOC that was launched in Fall of 2013, we were given permission by Coursera to crawl and study a small number of other courses. Our dataset consists of three courses: one social science course, “Accountable Talk: Conversation that works<sup>4</sup>”, offered in October 2013, which has 1,146 active users (active users refer to those who post at least one post in a course forum) and 5,107 forum posts; one literature course, “Fantasy and Science Fiction: the human mind, our modern world<sup>5</sup>”, offered in June 2013, which has 771 active users who have posted 6,520 posts in the course forum; one programming course, “Learn to Program: The Fundamentals<sup>6</sup>”, offered in August 2013, which has 3,590 active users and 24,963 forum posts. All three courses are officially seven weeks long. Each course has seven week specific subforums and a separate general subforum for more general discussion about the course. Our analysis is limited to behavior within the discussion forums.

## 4.3 Methods

### 4.3.1 Learner Motivation

Most of the recent research on learner motivation in MOOCs is based on surveys and relatively small samples of hand-coded user-stated goals or reasons for dropout [30, 37, 136]. Poellhuber et al. [136] find that user goals specified in the pre-course survey were the strongest predictors of later learning behaviors. Motivation is identified as an important determinant of engagement in MOOCs in the Milligan et al. [124] study. However, different courses design different enrollment motivational questionnaire items, which makes it difficult to generalize the conclusions from course to course. Another drawback is that learner motivation is volatile. In particular, distant learners can lose interest very fast even if they had been progressing well in the past [91]. It is important to monitor learner motivation and how it varies along the course weeks. We automatically measure learner motivation based on linguistic cues in the forum posts.

<sup>4</sup><https://www.coursera.org/course/accountabletalk>

<sup>5</sup><https://www.coursera.org/course/fantasysf>

<sup>6</sup><https://www.coursera.org/course/programming1>

### 4.3.2 Predicting Learner Motivation

The level of a student’s motivation strongly influences the intensity of the student’s participation in the course. Previous research has shown that it is possible to categorize learner motivation based on a students’ description of planned learning actions [58, 128]. The identified motivation categorization has a substantial relationship to both learning behavior and learning outcomes. But the lab-based experimental techniques used in this prior work are impractical for the ever-growing size of MOOCs. It is difficult for instructors to personally identify students who lack motivation based on their own personal inspection in MOOCs given the high student to instructor ratio. To overcome these challenges, we build machine learning models to automatically identify level of learner motivation based on posts to the course forum. We validate our measure in a domain general way by not only testing on data from the same course, but also by training on one course and then testing one other in order to uncover course independent motivation cues. The linguistic features that are predicative of learner motivation provide insights into what motivates the learners.

#### Creating the Human-Coded Dataset: MTurk

We used Amazon’s Mechanical Turk (MTurk) to make it practical to construct a reliable annotated corpus for developing our automated measure of student motivation. Amazon’s Mechanical Turk is an online marketplace for crowdsourcing. It allows requesters to post jobs and workers to choose jobs they would like to complete. Jobs are defined and paid in units of so-called Human Intelligence Tasks (HITs). Snow et al. [154] has shown that the combined judgments of a small number (about 5) of naive annotators on MTurk leads to ratings of texts that are very similar to those of experts. This applies for content such as the emotions expressed, the relative timing of events referred to in the text, word similarity, word sense disambiguation, and linguistic entailment or implication. As we show below, MTurk workers’ judgments of learner motivation are also similar to coders who are familiar with the course content.

We randomly sampled 514 posts from the Accountable Talk course forums and 534 posts from the Fantasy and Science Fiction course forums. The non-English posts were manually filtered out. In order to construct a hand-coded dataset for training machine learning models later, we employed MTurk workers to rate each message with the level of learner motivation towards the course the corresponding post had. We provided them with explicit definitions to use in making their judgment. For each post, the annotator had to indicate how motivated she perceived the post author to be towards the course by a 1-7 Likert scale ranging from “Extremely unmotivated” to “Extremely motivated”. Each request was labeled by six different annotators. We paid \$0.06 for rating each post. To encourage workers to take the numeric rating task seriously, we also asked them to highlight words and phrases in the post that provided evidence for their ratings. To further control the annotation quality, we required that all workers have a United States location and have 98% or more of their previous submissions accepted. We monitored the annotation job and manually filtered out annotators who submitted uniform or seemingly random annotations.

We define the motivation score of a post as the average of the six scores assigned by the annotators. The distributions of resulting motivation scores are shown in Figure 4.2. The following two examples from our final hand-coded dataset of the Accountable Talk class illustrate

the scale. One shows high motivation, and the other demonstrates low motivation. The example posts shown in this chapter are lightly disguised and shortened to protect user privacy.

- Learner Motivation = 7.0 (Extremely motivated)  
Referring to the last video on IRE impacts in our learning environments, I have to confess that I have been a victim of IRE and I can recall the silence followed by an exact and final received from a bright student.... Many ESL classes are like the cemetery of optional responses let alone engineering discussions. The Singing Man class is like a dream for many ESL teachers or even students if they have a chance to see the video! ...Lets practice this in our classrooms to share the feedback later!
- Learner Motivation = 1.0 (Extremely unmotivated)  
I have taken several coursera courses, and while I am willing to give every course a chance, I was underwhelmed by the presentation. I would strongly suggest you looking at other courses and ramping up the lectures. I'm sure the content is worthy, I am just not motivated to endure a bland presentation to get to it. All the best, XX.

### **Inter-annotator Agreement**

To evaluate the reliability of the annotations we calculate the intra-class correlation coefficient for the motivation annotation. Intra-class correlation [97] is appropriate to assess the consistency of quantitative measurements when all objects are not rated by the same judges. The intra-class correlation coefficient for learner motivation is 0.74 for the Accountable Talk class and 0.72 for the Fantasy and Science Fiction course.

To assess the validity of their ratings, we also had the workers code 30 Accountable Talk forum posts which had been previously coded by experts. The correlation of MTurkers' average ratings and the experts' average ratings was moderate ( $r = .74$ ) for level of learner motivation.

We acknowledge that the perception of motivation is highly subjective and annotators may have inconsistent scales. In an attempt to mitigate this risk, instead of using the raw motivational score from MTurk, for each course, we break the set of annotated posts into two balanced groups based on the motivation scores: "motivated" posts and "unmotivated" posts.

### **Linguistic Markers of Learner Motivation**

In this section, we work to find domain-independent motivation cues so that a machine learning model is able to capture motivation expressed in posts reliably across different MOOCs. Building on the literature of learner motivation, we design five linguistic features and describe them below. The features are binary indicators of whether certain words appeared in the post or not. Table 4.1 describes the distribution of the motivational markers in our Accountable Talk annotated dataset. We do not include the Fantasy and Science Fiction dataset in this analysis, because they will serve as a test domain dataset for our prediction task in the next section.

**Apply** words (Table 4.1, line 1): previous research on E-learning has found that motivation to learn can be expressed as the attention and effort required to complete a learning task and then apply the new material to the work site or life [62]. Actively relating learning to potential application is a sign of a motivated learner [125]. So we hypothesize that words that indicate application of new knowledge can be cues of learner motivation.

The Apply lexicon we use consists of words that are synonyms of “apply” or “use”: “apply”, “try”, “utilize”, “employ”, “practice”, “use”, “help”, “exploit” and “implement”.

**Need** words (Table 4.1, line 2) show the participant’s need, plan and goals: “hope”, “want”, “need”, “will”, “would like”, “plan”, “aim” and “goal”. Previous research has shown that learners could be encouraged to identify and articulate clear aims and goals for the course to increase motivation [114, 124].

**LIWC-cognitive** words (Table 4.1, line 3): The cognitive mechanism dictionary in LIWC [135] includes such terms as “thinking”, “realized”, “understand”, “insight” and “comprehend”.

**First person pronouns** (Table 4.1, line 4): using more first person pronouns may indicate the user can relate the discussion to self effectively.

**Positive words** (Table 4.1, line 5) from the sentiment lexicon [113] are also indicators of learner motivation. Learners with positive attitudes have been demonstrated to be more motivated in E-learning settings [125]. Note that negative words are not necessarily indicative of unmotivated posts, because an engaged learner may also post negative comments. This has also been reported in earlier work by Ramesh et al. [138].

The features we use here are mostly indicators of high user motivation. The features that are indicative of low user motivation do not appear as frequently as we expected from the literature. This may be partly due to the fact that students who post in the forum have higher learner motivation in general.

Feature	In Motivated post set	In Unmotivated post set
Apply**	57%	42%
Need**	54%	37%
LIWC -cognitive**	56%	38%
1st Person***	98%	86%
Positive***	91%	77%

Table 4.1: Features for predicting learner motivation. A binomial test is used to measure the feature distribution difference between the motivated and unmotivated post sets(\*\*:  $p < 0.01$ , \*\*\*:  $p < 0.001$ ).

## Experimental Setup

To evaluate the robustness and domain-independence of the analysis from the previous section, we set up our motivation prediction experiments on the two courses. We treat Accountable Talk as a “development domain” since we use it for developing and identifying linguistic features. Fantasy and Science Fiction is thus our “test domain” since it was not used for identifying the features.

For each post, we classify it as “motivated” or “unmotivated”. The amount of data from the two courses is balanced within each category. In particular, each category contains 257 posts from the Accountable Talk course and 267 posts for the Fantasy and Science Fiction course.

<b>Train Test</b>	<b>In-domain</b>		<b>Cross-domain</b>	
	Accountable Accountable	Fantasy Fantasy	Accountable Fantasy	Fantasy Accountable
<b>Unigram</b>	71.1%	64.0%	61.0%	61.3%
<b>Ling.</b>	65.2%	60.1%	61.4%	60.8%
<b>Unigram+Ling.</b>	72.3%	66.7%	63.3%	63.7%

Table 4.2: Accuracies of our three classifiers for the Accountable Talk course (Accountable) and the Fantasy and Science Fiction course (Fantasy), for in-domain and cross-domain settings. The random baseline performance is 50%.

We compare three different feature sets: a unigram feature representation as a baseline feature set, a linguistic classifier (Ling.) using only the linguistic features described in the previous section, and a combined feature set (Unigram+Ling.). We use logistic regression for our binary classification task. We employ liblinear [64] in Weka [197] to build the linear models. In order to prevent overfitting we use Ridge (L2) regularization.

### Motivation Prediction

We now show how our feature based analysis can be used in a machine learning model for automatically classifying forum posts according to learner motivation.

To ensure we capture the course-independent learner motivation markers, we evaluate the classifiers both in an in-domain setting, with a 10-fold cross validation, and in a cross-domain setting, where we train on one course’s data and test on the other (Table). For both our development (Accountable Talk) and our test (Fantasy and Science Fiction) domains, and in both the in-domain and cross-domain settings, the linguistic features give 1-3% absolute improvement over the unigram model.

The experiments in this section confirm that our theory-inspired features are indeed effective in practice, and generalize well to new domains. The bag-of-words model is hard to be applied to different course posts due to the different content of the courses. For example, many motivational posts in the Accountable Talk course discuss about teaching strategies. So words such as “student” and “classroom” have high feature weight in the model. This is not necessarily true for the other courses whose content has nothing to do with teaching.

In this section, we examine learner motivation where it can be perceived by a human. However, it is naive to assume that every forum post of a user can be regarded as a motivational statement. Many posts do not contain markers of learner motivation. In the next section, we measure the cognitive engagement level of a student based on her posts, which may be detectable more broadly.

### 4.3.3 Level of Cognitive Engagement

Level of cognitive engagement captures the attention and effort in interpreting, analyzing and reasoning about the course material that is visible in discussion posts [161]. Previous work uses manual content analysis to examine studentscognitive engagement in computer-mediated

communication (CMC) [63, 208]. In the MOOC forums, some of the posts are more descriptive of a particular scenario. Some of the posts contain more higher-order thinking, such as deeper interpretations of the course material. Whether the post is more descriptive or interpretive may reflect different levels of cognitive engagement of the post author. Recent work shows that level of language abstraction reflects level of cognitive inferences [21]. In this section, we measure the level of cognitive engagement of a MOOC user with the level of language abstraction of her forum posts.

## Measuring Level of Language Abstraction

Concrete words refer to things, events, and properties that we can perceive directly with our senses, such as “trees”, “walking”, and “red”. Abstract words refer to ideas and concepts that are distant from immediate perception, such as “sense”, “analysis”, and “disputable” [179].

Previous work measures level of language abstraction with Linguistic Inquiry and Word Count (LIWC) word categories [21, 135]. For a broader word coverage, we use the automatically generated abstractness dictionary from Turney et al. [179] which is publicly available. This dictionary contains 114,501 words. They automatically calculate a numerical rating of the degree of abstractness of a word on a scale from 0 (highly concrete) to 1 (highly abstract) based on generated feature vectors from the contexts the word has been found in.

The mean level of abstraction was computed for each post by adding the abstractness score of each word in the post and dividing that by the total number of words. The following are two example posts from the Accountable Talk course Week 2 subforum, one with high level of abstraction and one with low level of abstraction. Based on the abstraction dictionary, abstract words are in *italic* and concrete words are underlined.

- Level of abstraction = 0.85 (top 10%)

I *agree*. Probably *what* you just have to keep in mind is that you are there to help them *learn* by giving them opportunities to *REASON* out. In that case, you will not just *accept* the student’s *answer* but let them *explain* how they arrived towards it. Keep in mind to appreciate and *challenge* their *answers*.

- Level of abstraction = 0.13 (bottom 10%)

I teach science to gifted middle school students. The students learned to have conversations with me as a class and with the expert her wrote Chapter 8 of a text published in 2000. They are trying to design erosion control features for the building of a basketball court at the bottom of a hill in rainy Oregon.

We believe that level of language abstraction reflects the understanding that goes into using those abstract words when creating the post. In the Learn to Program course forums, many discussion threads are solving actual programming problems, which is very different from the other two courses where more subjective reflections of the course contents are shared. Higher level of language abstraction reflects the understanding of a broader problem. More concrete words are used when describing a particular bug a student encounters. Below are two examples.

- Level of abstraction = 0.65 (top 10%)

I have the same problems here. Make *sure* that your *variable* names match exactly. *Remember* that under-bars connect words together. I *know* something to do with the

read\_board(board\_file) function, but still need someone to *explain* more clearly.

- Level of abstraction = 0.30 (bottom 10%)  
>>> print(python, is)('python', 'is') >>> print('like', 'the', 'instructors', 'python') It leaves the 'quotes' and commas, when the instructor does the same type of print in the example she gets not parenthesis, quotes, or commas. Does anyone know why?

	Accountable Talk		Fantasy and Science Fiction		Learn to Program	
Control/Indep. Variable	p	Std. Err.	HR	p	HR	p
<b>Cohort1</b>	.68	.00	.82	.03	.81	.04
<b>PostCountByUser</b>	.86	.00	.90	.00	.76	.00
<b>AvgMotivation</b>	.58	.13	.82	.02	.84	.00
<b>AvgCogEngagement</b>	.94	.02	.92	.01	.53	.02

Table 4.3: Results of the survival analysis.

## 4.4 Validation Experiments

We use survival analysis (Chapter 3) to validate that participants with higher measured level of engagement will stay active in the forums longer, controlling for other forum behaviors such as how many posts the user contributes. We apply our linguistic measures described in Section 4 to quantify student engagement. We use the in-domain learner motivation classifiers with both linguistic and unigram features (Section 4.1.5) for the Accountable Talk class and the Fantasy and Science Fiction class. We use the classifier trained on the Accountable Talk dataset to assign motivated/unmotivated labels to the posts in the Learn to Program course.

### 4.4.1 Survival Model Design

For each of our three courses, we include all the active students, i.e. who contributed one or more posts to the course forums. We define the time intervals as student participation weeks. We considered the timestamp of the first post by each student as the starting date for that student's participation in the course discussion forums and the date of the last post as the end of participation unless it is the last course week.

*Dependent Variable:*

In our model, the dependent measure is **Dropout**, which is 1 on a student's last week of active participation unless it is the last course week (i.e. the seventh course week), and 0 on other weeks.

*Control Variables:*

**Cohort1:** This is a binary indicator that describes whether a user had ever posted in the first course week (1) or not (0). Members who join the course in earlier weeks are more likely than others to continue participating in discussion forums [200].

**PostCountByUser:** This is the number of messages a member posts in the forums in a week, which is a basic effort measure of activity level of a student.

**CommentCount:** This is the number of comments a user’s posts receive in the forums in a week. Since this variable is highly correlated with PostCountByUser ( $r > .70$  for all three courses). In order to avoid multicollinearity problems, we only include PostCountByUser in the final models.  
*Independent variables:*

**AvgMotivation** is the percentage of an individual’s posts in that week that are predicted as “motivated” using our model with both unigram and linguistic features (Section 4.1.4).

**AvgCogEngagement:** This variable measures the average abstractness score per post each week.

We note that AvgMotivation and AvgCogEngagement are not correlated with PostCountByUser ( $r < .20$  for all three courses). So they are orthogonal to the simpler measure of student engagement. AvgMotivation is not correlated with AvgAbstractness ( $r < .10$  for all three courses). Thus, it is acceptable to include these variables together in the same model.

## 4.4.2 Survival Model Results

Table 4.3 reports the estimates from the survival models for the control and independent variables entered into the survival regression.

Effects are reported in terms of the hazard ratio (HR), which is the effect of an explanatory variable on the risk or probability of participants drop out from the course forum. Because all the explanatory variables except Cohort1 have been standardized, the hazard rate here is the predicted change in the probability of dropout from the course forum for a unit increase in the predictor variable(i.e., Cohort1 changing from 0 to 1 or the continuous variable increasing by a standard deviation when all the other variables are at their mean levels).

Our variables show similar effects across our three courses (Table 4.3). Here we explain the results on the Accountable Talk course. The hazard ratio value for Cohort1 means that members survival in the group is 32% <sup>7</sup> higher for those who have posted in the first course week. Students’ activity level measure is correlated with student’s commitment. The hazard ratio for PostCountByUser indicates that survival rates are 14% <sup>8</sup> higher for those who posted a standard deviation more posts than average.

Controlling for when the participants started to post in the forum and the total number of posts published each week, both learner motivation and average level of abstraction significantly influenced the dropout rates in the same direction. Those whose posts expressed an average of one standard deviation more learner motivation (AvgMotivation) are 42% <sup>9</sup> more likely to continue posting in the course forum. Those whose posts have an average of one standard deviation higher cognitive engagement level (AvgCogEngagement) are 6% <sup>10</sup> more likely to continue posting in the course forum. AvgMotivation is relatively more predicative of user dropout than AvgCogEngagement for the Accountable Talk course and the Fantasy and Science Fiction course, while AvgCogEngagement is more predicative of user dropout in the Learn to Program course. This may be due to that in the Learn to Program course more technical problems are discussed and less posts contain motivation markers.

<sup>7</sup>32% = 100% - (100% \* 0.86)

<sup>8</sup>14% = 100% - (100% \* 0.86)

<sup>9</sup>42% = 100% - (100% \* 0.58)

<sup>10</sup>6% = 100% - (100% \* 0.94)



Figure 4.3- 4.5 illustrate these results graphically, showing three survival curves for each course. The solid curve shows survival with the number of posts, motivation, and cognitive engagement at their mean level. The top curve shows survival when the number of posts is at its mean level, and average learner motivation and level of cognitive engagement in the posts are both one standard deviation above the mean, and the bottom curve shows survival when the number of posts is at its mean, and the average expressed learner motivation and level of cognitive engagement in the posts are one standard deviation below the average.

### 4.4.3 Implications

In contrast to regular courses where students engage with class materials in a structured and monitored way, and instructors directly observe student behavior and provide feedback, in MOOCs, it is important to target the limited instructor's attention to students who need it most [138]. The automated linguistic models designed in this chapter can help monitor MOOC user engagement from forum posts. By identifying students who are likely to end up not completing the class before it is too late, we can perform targeted interventions (e.g., sending encouraging emails, posting reminders, allocating limited tutoring resources, etc.) to try to improve the engagement of these students. For example, our motivation prediction model could be used to improve targeting of limited instructor's attention to users who are motivated in general but are experiencing a temporary lack of motivation that might threaten their continued participation, in particular, those who have shown serious intention of finishing the course by joining the discussion forums. One possible intervention that can be based on this type of analysis might suggest instructors reply to those with recent motivation level lower than it has been in the past. This may help students get past a difficult part of the course. We can also recommend reading the highly motivated posts to the other users, which may serve as an inspiration.

Based on the predictive engagement markers, we see it is important for the students to be able to apply new knowledge and engage in deeper thinking. Discussion facilitation can influence levels of cognitive engagement [46]. The instructor can encourage learners to reflect on what and how learning addressed needs. Work on automated facilitation from the Computer Supported Collaborative Learning (CSCL) literature might be able to be adapted to the MOOC context to make this feasible [3].

## 4.5 Chapter Discussions

The goal of this thesis is to improve students' commitment and collaboration product in the MOOC context. To this end, we investigate which process measures are related to these outcomes. In xMOOCs, we found that students' activity level and engagement cues extracted from students' forum posts were correlated with student commitment. We identify two new measures that quantify engagement and validate the measures on three Coursera courses with diverse content. We automatically identify the extent to which posts in course forums express learner motivation and cognitive engagement. The survival analysis results validate that the more motivation the learner expresses, the lower the risk of dropout. Similarly, the more personal interpretation a participant shows in her posts, the lower the rate of student dropout from the course forums.

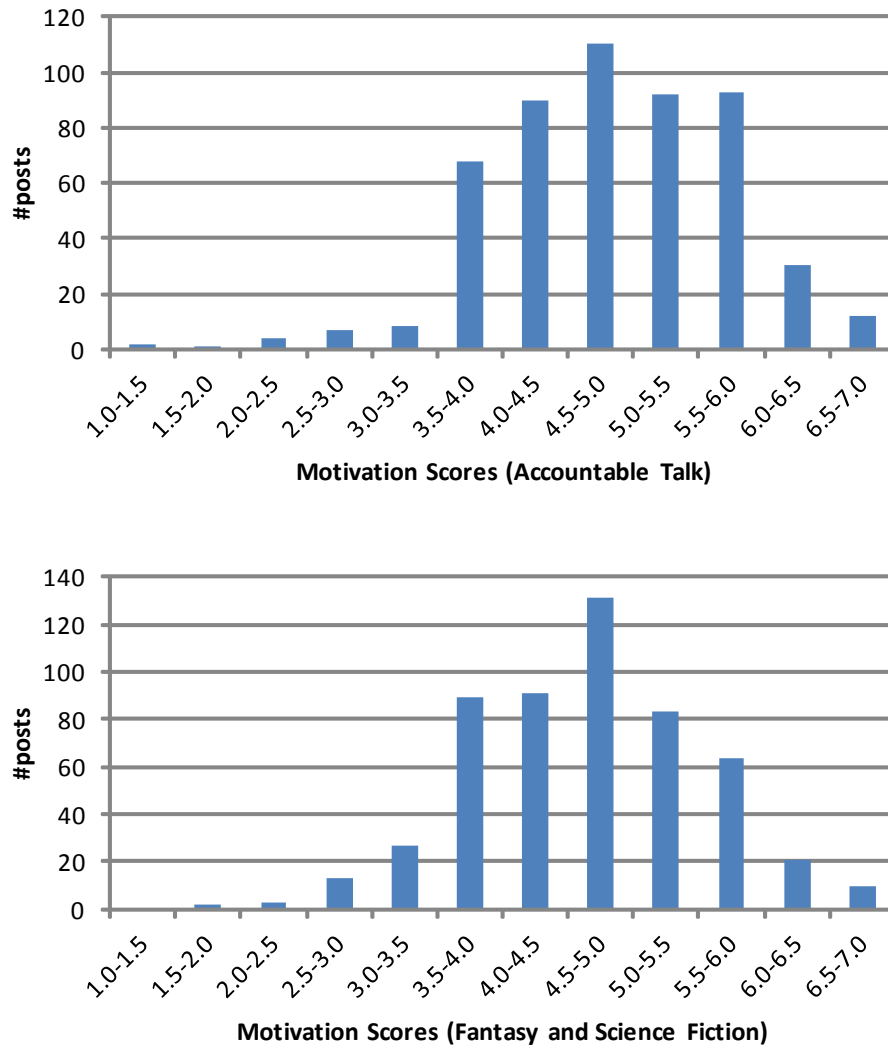


Figure 4.2: Annotated motivation score distribution.

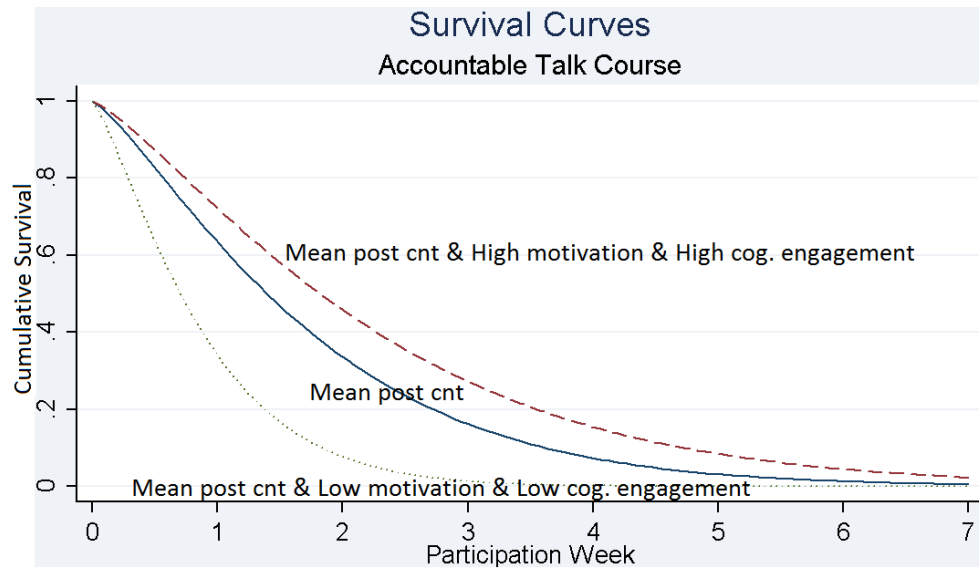


Figure 4.3: Survival curves for students with different levels of engagement in the Accountable Talk course.

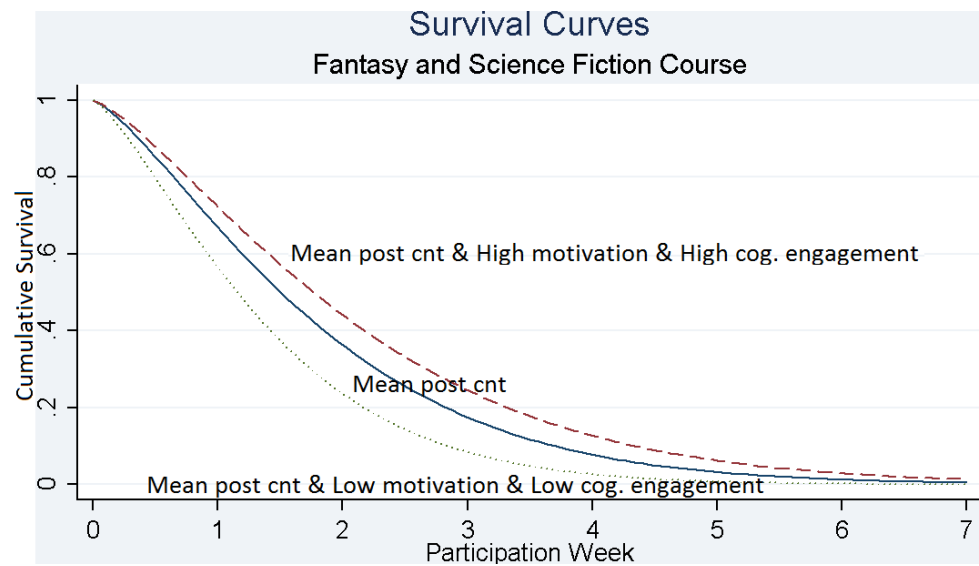


Figure 4.4: Survival curves for students with different levels of engagement in the Fantasy and Science Fiction course.

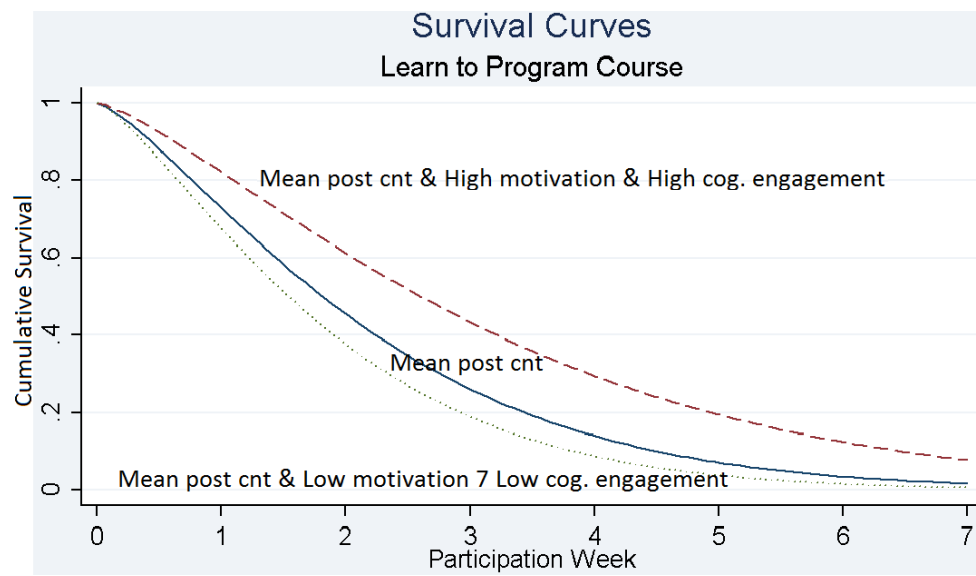


Figure 4.5: Survival curves for students with different levels of engagement in the Learn to Program course.

## Chapter 5

# Virtual Teams in Massive Open Online Courses: Study 2

<sup>1</sup> In Study 1, we identified what factors were correlated with student's commitment in xMOOCs. Team-based learning, as realized in MOOCs have many factors that may positively or negatively impact students commitment and collaboration product. In order to support team-based MOOCs, in this chapter, we study what process measures are correlated with commitment and team collaboration product in team-based MOOCs.

Most current MOOCs have scant affordances for social interaction, and arguably, that social interaction is not a major part of the participation experience of the majority of participants. We first explore properties of virtual team's social interaction in a small sample of xMOOCs where groups are spontaneously formed and more social-focused NovoEd MOOCs where team-based learning is an integral part of the course design. We show that without affordances to sustain group engagement, students who joined study groups in xMOOC forum does not have a higher commitment to the course.

Less popular, emerging platforms like NovoEd<sup>2</sup> are designed differently, with team-based learning or social interaction at center stage. In NovoEd MOOCs where students work together in virtual teams, similar to our findings in Study 1, activity level and engagement cues extracted from team communication are both correlated with team collaboration product quality. Team leader's leadership behaviours, which is one type of collaboration communication support, play a critical role in supporting team collaboration discussions. Figure 5.1 presents our hypotheses in this chapter.

MOOC	#Forum Users	#Users in Study Group(%)	#Groups	#Course Weeks
Financial Planning	5,305	1,294(24%)	121	7
Virtual Instruction	1,641	278(17%)	22	5
Algebra	2,125	126(6%)	23	10

Table 5.1: Statistics of the three Coursera MOOCs.

<sup>1</sup>The contents of this chapter are modified from a published paper: [195]

<sup>2</sup><https://novoed.com/>

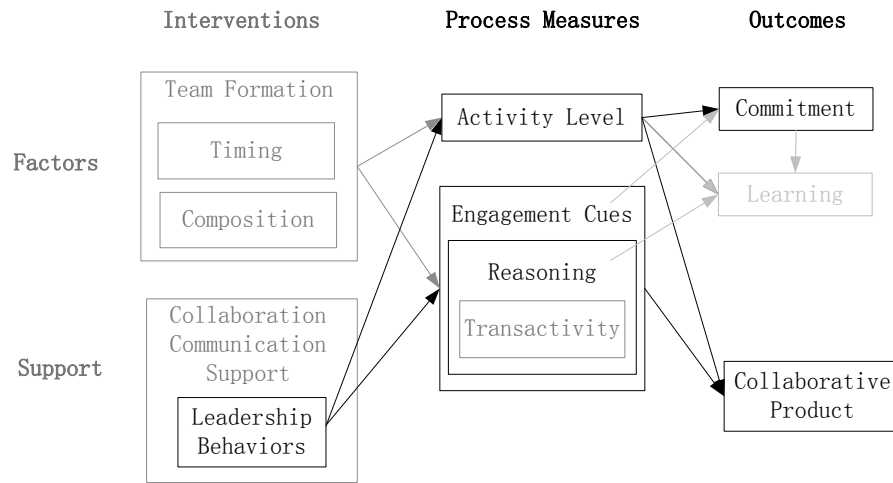


Figure 5.1: Chapter hypothesis.

NovoEd	#Registered Users	#Users successfully joined a group	# Teams	# Course Weeks
Elementary	2,817	262	101	12
Secondary	1,924	161	76	12

Table 5.2: Statistics two NovoEd MOOCs.

## 5.1 MOOC Datasets

Our xMOOC dataset consists of three Coursera<sup>4</sup> MOOCs, one about virtual instruction, one that is an algebra course, and a third that is about personal financial planning. These MOOCs were offered in 2013. The statistics are shown in Table 5.1.

Our NovoEd dataset consists of two NovoEd MOOCs. Both courses are teacher professional development courses about Constructive Classroom Conversations, in elementary and secondary education. They were offered simultaneously in 2014. The statistics are shown in Table 5.2.

<sup>4</sup><https://www.coursera.org/>

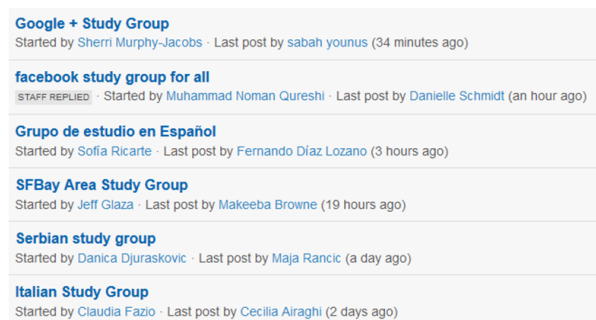


Figure 5.2: Screen shot of a study group subforum in a Coursera MOOC.

## 5.2 Study Groups in xMOOC Discussion Forums

To understand students' social need in state-of-the-art xMOOCs, we studied the ad hoc study groups formed in xMOOC course discussion forums. We show that many students indicate high interest in social learning in the study group subforums. However, there is little sustained communication in these study groups. Our longitudinal analysis results suggest that without affordances to sustain group engagement, groups quickly lose momentum, and therefore members in the group does not show higher retention rate.

In this section, we briefly introduce study groups as they exist in xMOOCs. In the default Coursera MOOC discussion forums, there is a "Study Groups" subforum where students form study groups through asynchronous threaded discussion. The study groups can be organized around social media sites, languages, districts or countries (Figure 5.2), and generally become associated with a single thread. Across our three xMOOCs, around 6-24% forum users post in the study group forums, indicating need for social learning (Table 5.1). More than 80% of students who post in some study group thread post there only one or two times. They usually just introduce themselves without further interacting with the group members in the study group thread, then the thread dies. More than 75% of the posts in the study group thread are posted in the first two weeks of the course.

### 5.2.1 Effects of joining study groups in xMOOCs

In this section, we examine the effects of joining study groups on students' drop out rate in xMOOCs with survival modeling. Survival models can be regarded as a type of regression model, which captures influences on time-related outcomes, such as whether or when an event occurs. In our case, we are investigating the influence of participation in study groups on the time when a course participant drops out of the course forum. More specifically, our goal is to understand whether our automatic measures of student engagement can predict her length of participation in the course forum. Survival analysis is known to provide less biased estimates than simpler techniques (e.g., standard least squares linear regression) that do not take into account the potentially truncated nature of time-to-event data (e.g., users who had not yet left the community at the time of the analysis but might at some point subsequently). From a more technical perspective, a survival model is a form of proportional odds logistic regression, where a prediction about the likelihood of a failure occurring is made at each time point based on the presence of some set of predictors. The estimated weights on the predictors are referred to as hazard ratios. The hazard ratio of a predictor indicates how the relative likelihood of the failure occurring increases or decreases with an increase or decrease in the associated predictor. We use the statistical software package Stata<sup>5</sup>. We assume a Weibull distribution of survival times, which is generally appropriate for modeling survival[151].

For each of our three courses, we include all the students, who contributed one or more posts to the course forums in the survival analysis. We define the time intervals as student participation weeks. We considered the timestamp of the first post by each student as the starting date for that student's participation in the course discussion forums and the date of the last post as the end of

<sup>5</sup><http://www.stata.com/>

participation unless it is the last course week.

*Dependent Variable:*

In our model, the dependent measure is **Dropout**, which is 1 on a student’s last week of active participation unless it is the last course week, and 0 on other weeks.

*Control Variables:*

**Cohort1** is a binary indicator that describes whether a user had ever posted in the first course week (1) or not (0). Members who join the course in earlier weeks are more likely than others to continue participating in discussion forums[193].

**PostCountByUser** is the number of messages a member posts in the forums in a week, which is a basic effort measure of engagement of a student.

*Independent variable:*

**JoinedStudyGroup** is a binary indicator that describes whether a user had ever posted in a study group thread(1) or not(0). We assume that students who have posted in a study group thread are members in a study group.

Variable	Financial Planning			Virtual Instruction			Algebra	
	Haz. Ratio	Std. Err.	$P >  z $	Haz. Ratio	Std. Err.	$P >  z $	Haz. Ratio	Std. Err.
PostCountByUser	0.69	0.035	0.000	0.65	0.024	0.000	0.74	0.026
Cohort1	0.74	0.021	0.000	0.66	0.067	0.000	0.74	0.056
JoinedStudyGroup	1.16	0.123	0.640	0.368	0.027	0.810	0.82	0.081

Table 5.3: Survival Analysis Results.

## Survival Analysis Results

Table 5.3 reports the estimates from the survival models for the control and the independent variable entered into the survival model. JoinedStudyGroup makes no significant prediction about student survival in this course. Across all the three Coursera MOOCs, the results indicate that students who join study groups in MOOC do not stay significantly longer in the course than the other students who posted at least once in the forum in any other area of the forum.

Current xMOOC users tend to join study groups in the course forums especially in the first several weeks of the course. However, we did not observe the benefit or influence of the study groups on commitment to the course. Since most students only post in the study group thread only once or twice in the beginning of the course, when students come back to the study group thread to ask for help later in the course, they cannot get the support from their teammates as the thread has been inactive. The substantial amount of students posting in the study group threads demonstrate students’ intention to “get through the course together” and the need for team based learning interventions. As is widely criticized, the current xMOOC platforms fail to provide the social infrastructure to support sustained communication in study groups.



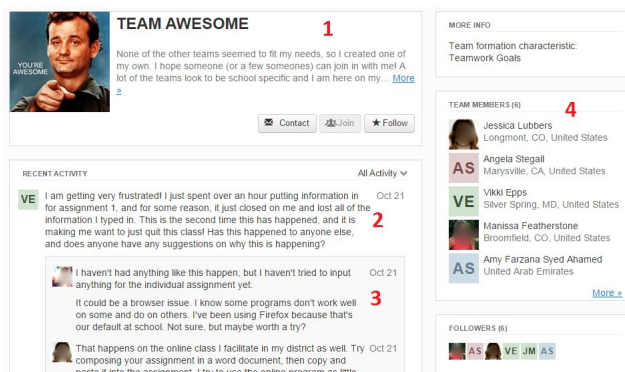


Figure 5.3: Homepage of a NovoEd team. 1: Team name, logo and description. 2: A team blog. 3: Blog comments. 4: Team membership roster.

## 5.3 Virtual Team in NovoEd MOOCs

In this section, we describe the experience of virtual teams in NovoEd, where team based learning is in the central of the course design. Then we demonstrate the teammate's major influence on a student's dropout through survival modeling.

### 5.3.1 The Nature of NovoEd Teams

Students in a NovoEd MOOC have to initiate or join a team in the beginning of the course. The student who creates the team will be the team leader. The homepage of a NovoEd team displays the stream of blog posts, events, files and other content shared with the group as well as the activate members (Figure 5.3). All the teams are public, with their content visible to anyone in the current NovoEd MOOC. Students can also communicate through private messages. In our NovoEd MOOCs, 35% of students posted at least one team blog or blog comment. 68% of students sent at least one private message. Only 4% of students posted in the general course discussion forums. Most students only interact with their teammates and TAs.

When a group is created, its founder chooses its name and optionally provides additional description and a team logo to represent the group. The founder of a team automatically becomes the team leader. The leader can also select classmates based on their profiles and send them invitation messages to join the team. The invitation message contains a link to join the group. Subsequently, new members may request to join and receive approval from the team leader. Only team leader can add a member to the team or delete the team.

Throughout the course, team work is a central part of the learning experience. In our NovoEd courses, instructors assign small tasks ("Housekeeping tasks", which will not be graded) such as "introduce yourself to the team" early on in the course. They also encourage students to collaborate with team members on non-collective assignments as well. Individual performance in a group is peer rated so as to encourage participation and contribution. Collaborations among students are centered on the final team assignment, which accounts for 20% of the final score.

Compared to study groups in an xMOOC, the virtual team leader has a much bigger impact on its members in a NovoEd MOOC. In our dataset, when the team leader drops out, 71 out of

84 teams did not submit the final team assignment.

### 5.3.2 Effects of teammate dropout

In this section, We use survival analysis to validate that when team leader or teammates drop out, the team member is more prone to dropout, but the effect of losing the team leader is stronger.

*Dependent Variable:*

We consider a student to drop out if the current week is his/her last week of active participation unless it is the last course week (i.e. the twelfth course week).

*Control Variables:*

**Activity:** total number of activities (team blogs, blog comments or messages) a student participated in that week, which is a basic effort measure of engagement of a student.

**GroupActivity:** total number of activities the whole team participated in that course week. Since it is correlated with Activity ( $r > .50$ ). In order to avoid multicollinearity problems, we only include Activity in the final survival models.

*Independent variables:*

**TeamLeaderDropout:** 1 if the team leader dropped out in previous weeks, 0 otherwise.

**TeammateDropout:** 1 if at least one of the teammates (besides the team leader) dropped out in the current week, 0 otherwise.

**RandomTeammateDropout:** to control for the effect that students in the MOOC may dropout at around the same time[201], we randomly sample a group of classmates for each student(“random team”), analogous to “nominal groups” in studies of process losses in the group work literature. The random team is the same size as the student’s team. 1 if at least one of the students in the random team dropped out in the current week. 0 otherwise.

Variable	Model 1			Model 2		
	Haz. Ratio	Std. Err.	P >  z	Haz. Ratio	Std. Err.	P >  z
Activity	0.654	0.091	0.002	0.637	0.085	0.001
TeamLeaderDropout	3.420	0.738	0.000			
TeammateDropout	2.338	0.576	0.001			
RandomTeammateDropout				1.019	0.273	0.945

Table 5.4: Survival analysis results.

### Survival Analysis Results

Since the survival analysis results are similar for the two NovoEd courses, we include all the 423 users who successfully joined a team for the two courses. Table 5.4 reports the estimates from the survival models for the control and independent variables entered into the survival regression. Effects are reported in terms of the hazard ratio, which is the effect of an explanatory variable on the risk or probability of participants drop out from the course forum. Students are 35% times less likely to dropout if they have one standard deviation more activities. **Model 1** reports when controlling for team activity, a student is more than three times more likely to drop out if

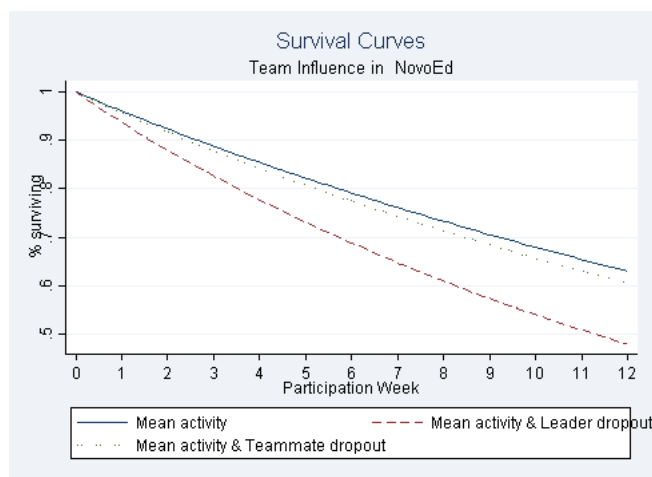


Figure 5.4: Survival plots illustrating team influence in NovoEd.

his/her team leader has dropped out. A student is more than two times more likely to drop out if his/her team leader drops out that week. Figure 5.4 illustrates our results graphically. The solid curve shows survival with user Activity at its mean level. The curve in the middle shows survival when user Activity is at its mean level and at least one of the teammates drop out in the current week. The curve in the middle shows survival when user Activity is at its mean level and team leader drops out in the current week. **Model 2** shows when controlling for team activity, “random teammates”(randomly sampled classmates) dropout does not have a significant effect on a student’s dropout.

If we compare the experience of social interaction between a typical xMOOC where groups are ad hoc course addons and a NovoEd course where they are an integral part of the course design, we see that making the teamwork a central part of the design of the curriculum encourages students to make that social interaction a priority. With teammates’ activities more visible, students are more influenced by their teammates’ activities. Teammates especially team leader behavior has a big impact on a student’s engagement in the course.

## 5.4 Predictive Factors of Team Performance

In the previous section, we demonstrate the importance of a team leader’s leadership behaviors. Not all team leaders are equally successful in their role. In this section, we model the relative importance of different team leaders behaviors we are able to detect. Analysis of what distinguishes successful and non-successful teams shows the important, central role of a team leader who coordinates team collaboration and discussion. In this section, we describe how much each behavior contributes to team success.

Despite the importance of team performance and functioning in many workplaces, relatively little is known about effective and unobtrusive indicators of team performance and processes [25]. Drawing from previous research we design three latent factors we referred to as Team Activity, Communication Language and Leadership Behavior, that we hypothesis are predicative

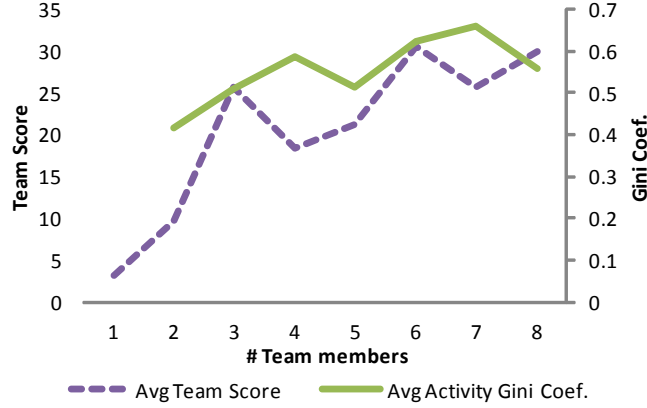


Figure 5.5: Statistics of team size, average team score and average overall activity Gini coefficient.

of team performance. In this section, we further formalize these latent factors by specifying associated sets of observed variables that will ultimately enable us to evaluate our conceptual model.

The performance measure we use is the final **Team Score**, which was based on the quality of the final team project submission. The score can take value of 0, 20 or 40.

### 5.4.1 Team Activity

In this section, we describe the variables that account for variation in the level of activity across teams.

The control variables for volume of communication-related factors: **MemberCnt** is the number of members in the team. Group size is an important predictor of group success. **BlogCnt** is the number of team blogs that are posted. **BlogCommentCnt** is the number of team blog comments that are made. **MessageCnt** is the number of messages that are sent among the team. **Equal Participation** is the degree to which all group members are equally involved, can enable everyone in the group to benefit from the team. In consistent with Borge et al. [24], we measure equal participation with Gini coefficient. In our dataset, unequal participation is severer in larger teams and actually indicates high quality work[93].

### 5.4.2 Communication Language

Groups that work well together typically exchange more knowledge and establish good social relationships, which is reflected in the way that they use words[120]. We first include three measures that are predictive of team performance in previous work[169]. **Positivity** is the percentage of posts that contain at least one positive word in LIWC dictionary[135]. It measures the degree to which group members encourage one another by offering supportive behaviors that enhance interpersonal relationships and motivate individuals to work harder[110]. Since negative words tend not to be used, they are also not significant predictors of team performance, and thus we did

not include a measure of them in the final model. **Information exchange** is commonly measured by word count[149, 169]. **Engagement** is the degree to which group members are paying attention and connecting with each other. It can enhance group cohesion. Engagement is reflected in the degree to which group members tend to converge in the way they talk, which is called Linguistic Style Matching (LSM)[86, 129]. We measure LSM using asynchronous communication (messages) as the input[127]. We excluded automated emails generated from NovoEd. For consistency with prior work, we employed the nine LIWC-derived categories[129]. Our nine markers are thus: articles, auxiliary verbs, negations, conjunctions, high-frequency adverbs, impersonal pronouns, personal pronouns, prepositions, and quantifiers. For each of the nine categories  $c$ , the percentage of an individual  $n$ 's total words ( $p_{c,n}$ ) was calculated, as well as the percentage of the group's total words ( $p_{Gc}$ ). This allows the calculation of an individual's similarity against the group, per word category, as

$$LSM_{c,n} = 1 - \frac{|p_{c,n} - p_{Gc}|}{p_{c,n} + p_{Gc}}$$

The group  $G$ 's average for category  $c$  is the average of the individual scores:

$$LSM_{G,c} = \frac{\sum_{n \in G} LSM_{c,n}}{|G|}$$

And the team LSM is the average across the nine categories:

$$LSM_G = \frac{\sum_{c=1,2,\dots,9} LSM_{G,c}}{9}$$

We design three new language indicators. **1st Person Pronouns** is the proportion of first-person singular pronouns can suggest self-focused topics. **Tool** is the percentage of blogs or messages that mention a communication/collaboration tool, such as Google Doc/Hangout, Skype or Powerpoint. **Politeness** is the average of the automatically predicted politeness scores of the blog posts, comments and messages[49]. This factor captures how polite or friendly the team communication is, based on features like mentions of “thanks”, “please”, etc.

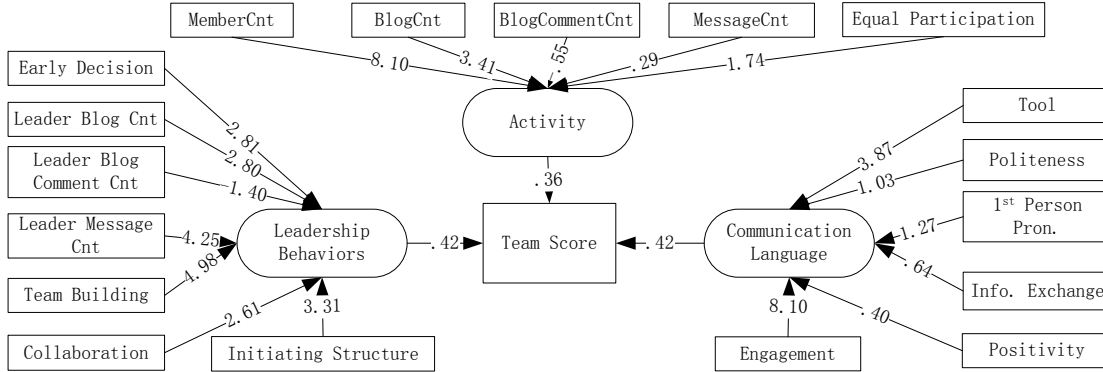


Figure 5.6: Structural equation model with maximum likelihood estimates (standardized). All the significance level  $p < 0.001$ .

### 5.4.3 Leadership Behaviors

Leaders are important for the smooth functioning of teams, but their effect on team performance is less well understood[60]. We identified three main leadership behaviors based on the messages

sent by team leaders (Table 5.5). These behaviors are mainly coordinating virtual team collaboration and discussion. A message will be coded as “other” if it contains none of the behaviors listed in Table 5.5. 30 messages sent by team leaders are randomly sampled and then coded by two experts. Inter-rater reliability was Kappa = .76, indicating substantial inter-rater reliability. Then one of the experts code all the 855 leader messages in the two NovoEd MOOCs into these three categories.

Type	Behaviors	Example Team Leader Message
Team Building	Invite or accept users to join the group	<i>Lauren, We would love to have you. Jill and I are both ESL specialists in Boston.</i>
Initiating Structure	Initiate a task or assign subtask to a team member	<i>Housekeeping Task #3 is optional but below are the questions I can summarize and submit for our team.</i>
Collaboration	Collaborate with teammates, provide help or feedback	<i>I figured out how to use the Google Docs. Let's use it to share our lesson plans.</i>

Table 5.5: Three different types of leadership behaviors.

Thus we design three variables to characterize team leader's coordination behavior. **Team Building** is the number of Team Building messages sent by the team leader. Team Building behavior is critical for NovoEd teams since only the team leader can add new members to the team. **Initiating Structure** is the number of Initiating Structure messages sent by the team leader. Initiating structure is a typical task-oriented leadership behavior. Previous research has demonstrated that in comparison to relational-oriented or passive leader behavior, task-oriented leader behavior makes a stronger prediction about task performance[56]. Initiating Structure is also an important leadership behavior in virtual teams. When the team is initially formed, usually the team leader needs to break the ice to start to introduce him/herself. When assignment deadlines are coming up, usually team leaders need to call that to the attention of teammembers to start working on it. **Collaboration** is the number of Collaboration messages sent by the team leader. Different from previous research where the team leader mainly focuses on managing the team, leaders in virtual teams usually take on important tasks in the team project as well, which is reflected in Collaboration behaviors.

Another variable, **Early Decision**, is designed to control for a team leader's early decisions. In addition to choosing a name, which all groups must have, leaders can optionally write a description that elaborates on the purpose of the group, and 82% did so. They can upload a photo as team logo that will appear with the group's name on its team page and in the course group lists, and 54% did so. The team leader should add a team avatar to attract more team members. Since these are optional, we construct the variable 0 if the team does none. 0.5 if one is done. 1 if both are done.

We include three control variables to control for team leader's amount of activity: **LeaderBlogCnt** is the number of team blogs that are posted that course week. **LeaderBlogCommentCnt** is the number of team blogs comments that are made that course week. **LeaderMessageCnt** is the number of messages that are sent among the team during that course week.

#### 5.4.4 Results

In Section 5.4.1 - 5.4.3, we described three latent factors we hypothesize are important in distinguishing successful and unsuccessful teams along with sets of associated observed variables. In this section, we validate the influence of each latent factor on team performance using a generalized Structural Equation Model (SEM) in Stata. Experiments are conducted on all the 177 NovoEd MOOC teams.

Figure 5.6 shows the influence of each observed variable on its corresponding latent variable, and in turn the latent variable on team performance. The weights on each directed edge represent the standard estimated parameter for measuring the influence. Based on Figure 5.6, firstly, Leader Behaviors and Communication Language contribute similarly to team performance, with a standard estimated parameter of 0.42. Among the three leadership behaviors, Team Building is the strongest predictor of team performance. Since the number of Team Building messages significantly correlates with the total number of members in the team ( $r = 0.62^{***}$ ), consistent with prior work[102], team leader who recruits a larger group can increase the group's chances of success.

Among the Communication Language indicators, Engagement and the use of communication Tool are the most predictive. This demonstrates the importance of communication for virtual team performance. Utilizing various communication tool is an indicator of team communication. Different from previous work on longer-duration project teams[127], our results support that linguistic style matching is correlated with team performance. Team Activity is a comparatively weaker predictor of team performance. This is partly due to that some successful teams communicate through emails and have smaller amount of activities in the NovoEd site.

### 5.5 Chapter Discussions

In this chapter, we examined virtual teams in typical xMOOCs where groups are ad hoc course add-ons and NovoEd MOOCs where they are an integral part of the course design. The study groups in xMOOCs demonstrate high interest in social learning. The analysis results indicate that without the social environment that is conducive to sustained group engagement, members do not have the opportunity to benefit from these study groups. Students in NovoEd virtual teams have a more collaborative MOOC experience. Despite the fact that instructors and TAs try hard to support these team, half of the teams in our two NovoEd MOOCs are not able to submit a team project at the end of the course. Our longitudinal survival analysis shows that leader's dropout is correlated with much higher team member dropout. This result suggests that leaders can also be a potential point of failure in these virtual teams since they concentrate too much responsibility in themselves, e.g. only the leader can add new members to the team. Given the generally high dropout rate in MOOCs and the importance of the leadership behaviors, these findings suggest that to support virtual teams in NovoEd teams, we need to better support or manage collaborative communication.

Study 2 results indicate that we should either support the collaboration communication or improve group formation to improve team collaboration product quality. Since many teams do not have any activities after the team is formed, indicating that the current self selection based

team formation might have some problems. Thus we believe supporting group formation has more potential value. We hypothesis that we can support efficient team formation through asynchronous online deliberation. Besides this, our dataset in Study 2 has little to help us answer questions about how the teams should be formed and how to support that team formation process in MOOCs. In Study 3 and 4, we address this question by testing our team formation hypothesis in a crowdsourcing environment. Since leaders' role is mainly to facilitate collaboration communication, we will further investigate incorporating discussion facilitation agent in team collaboration in Study 5.



## Chapter 6

# Online Team Formation through a Deliberative Process in Crowdsourcing Environments: Study 3 and 4

In Study 1 and 2, I studied the factors (e.g. engagement cues, role taking behavior) that have correlation relationship with the outcome measures (e.g. retention rate, performance). Through these analysis, we conclude that a potential area or impact is to support the online team formation process. The experiments in this chapter will serve as our pilot study for the team formation intervention studies in Study 6 and 7. Rather than trying out untested designs on real live courses, we will first prototype and test our group formation procedure using a crowd-sourcing service, Amazon Mechanical Turk (MTurk), on a collaborative task. Previous work demonstrates that it is challenging to form groups either through self-selection or designation while ensuring high-quality and satisfactory collaboration [102, 164, 206]. We look to see if the positive group learning results that have been found in in-person classes and Computer-Supported Collaborative Learning (CSCL) contexts can be successfully replicated and extended in a more impersonal online context.

### 6.1 Introduction

One potential danger in small group formation is that the small groups lose contact with intellectual resources that the broader community provides. This is especially true where participants join teams very early in their engagement with the community and then focus most of their social energy on their group, as is the norm in team-based MOOCs [141]. Building on research in the learning sciences, we design a deliberation-based team formation strategy where students participate in a deliberation forum before working within a team on a collaborative learning assignment. The key idea of our team formation method is students should have the opportunity to interact meaningfully with the community before assignment into teams. That discussion provides students with a wealth of insight into alternative task-relevant perspectives to take with them into the collaboration [29]. To understand the effects of a deliberation approach to team formation, we ran a series of studies in MTurk [41]. The team formation process begins with

individual work where participants learned about one of four different energy types, then participated in a open discussion, and then worked with three other teammates (using a Jigsaw learning configuration [11]) to solve an energy-related challenge. We assessed team success based on the quality of the produced team proposal.

Study 3 tests the extent to which teams that engage in a large community forum deliberation process prior to team formation achieve better group task outcomes than teams that instead perform an analogous deliberation within their team. Simply stated, our first research question is:

*RQ1. Will exposure to large community discussions lead to more successful small group collaborations compared to teams that formed before community discussions compared to teams that formed after community discussions?*

Second, A group formation method has been identified as an important factor for enhancing small group outcomes [85]. Most of the previous research on small group formation focuses on issues related to group composition, for example, assigning participants based on expertise or experience [76]. This information may not be readily available in MOOCs. To address the disadvantage that small groups formed with limited prior communication might lack the synergy to work effectively together, our research further explores the hypothesis that participants' interactions during community deliberations may provides evidence of which students would work well together [89]. Our team formation matches students with whom they have display successful team processes, i.e., where transactive discussion has been exchanged during the deliberation. A transactive discussion is one where participants "elaborate, build upon, question or argue against previously presented ideas" [20]. This concept has also been referred to as intersubjective meaning making [165], productive agency [147], argumentative knowledge construction [190], idea co-construction [77]. It has long been established that transactive discussion is an important process that reflects good social dynamics in a group [172] and results in collaborative knowledge integration [51]. We leverages learning sciences findings that students who show signs of a good collaborative process form more effective teams [101, 144, 145]. In Study 4, we look for evidence of participants transactively reasoning with each other during community-wide deliberation and use it as input into a team formation algorithm. Simply stated, our second research question is:

*RQ2. Can evidence of transactive discussions during deliberation inform the formation of more successful teams compared to randomly formed teams?*

Figure 6.1 presents our hypotheses in this chapter.

To enable us to test the causal connection between variables in order to identify principles that later we will test in an actual MOOC, we validate our hypotheses underlying potential team formation strategies using a collaborative proposal-writing task hosted on the Amazon Mechanical Turk crowdsourcing platform (MTurk). The crowdsourcing environment offers a context in which we are able to study how the process of team formation influences the effectiveness with which the teams accomplish a short term task. While crowd workers likely have some different motivations from MOOC participants, the remote individual work setting without peer contact resembles the experience in many online settings. Since affordances for forum deliberation are routinely available in most online communities, we explore a deliberation-based team formation strategy that forms worker teams based on their interaction in a discussion forum where they engage in a discussion on a topic related to the upcoming team task.

Our results show that teams that form after community deliberation have better team perfor-

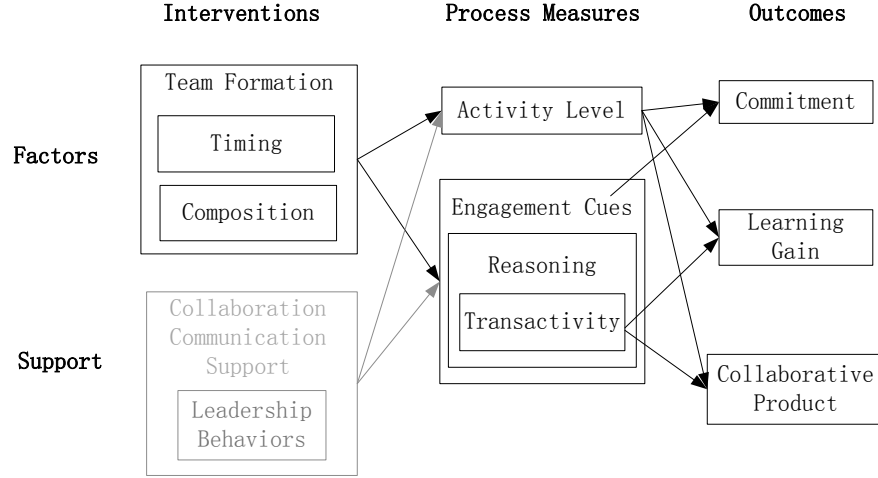


Figure 6.1: Chapter hypothesis.

mance than those that form before deliberation. Teams that have more transactive communication during deliberation have better team performance. Based on these positive results we design a community-level deliberative process for team formation that contributes towards effective team formation both in terms of preparing team members for effective team engagement, and for enabling effective team assignment.

## 6.2 Method

### 6.2.1 Experimental Paradigm

#### Collaboration Task Description

We designed a highly-interdependent collaboration task that requires negotiation in order to create a context in which effective group collaboration would be necessary for group success. In particular, we used a Jigsaw paradigm, which has been demonstrated as an effective way to achieve a positive group composition and is associated with positive group outcomes [11]. In a Jigsaw task, each participant is given a portion of the knowledge or resources needed to solve the problem, but no one has enough to complete the task alone.

Specifically, we designed a constraint satisfaction proposal writing task where the constraints came from multiple different perspectives, each of which were represented by a different team member. The goal was to require each team member to represent their own assigned perspective, but to consider how it related to the perspectives of others within the group. However, since a short task duration was required in order for the task to be feasible in the MTurk environment, it was necessary to keep the complexity of the task relatively low. In order to meet these requirements, the selected task asked teams to consider municipal energy plan alternatives that involved combinations of four energy sources (coal, wind, nuclear and hydro power) each paired with specific advantages and disadvantages. Following the Jigsaw paradigm, each member of the team was given special knowledge of one of the four energy sources, and was instructed to

In this final step, you will work together with other Turkers to recommend a way of distributing resources across energy types for the administration of City B. City B requires 12,000,000 MWh electricity a year from four types of energy sources: coal power, wind power, nuclear power and hydro power. We have provided 4 different plans to choose from, each of which emphasizes one energy source as primary. Specifically the plans describe how much energy should be generated from each of the four energy sources, listed in the table below. Your team needs to negotiate which plan is the best way of meeting your assigned goals, given the city's requirements and information below.

City B's requirements and information:

1. City B has a tight yearly energy budget of \$900,000K. Coal power costs \$40/MWh. Nuclear power costs \$100/MWh. Wind power costs \$70/MWh. Hydro power costs \$100/MWh.
2. The city is concerned with chemical waste. If the main energy source releases toxic chemical waste, there is a waste disposal cost of \$2/MWh.
3. The city is a famous tourist city for its natural bird and fish habitats.
4. The city is trying to reduce greenhouse gas emissions. If the main energy source release greenhouse gases, there will be a "Carbon tax" of \$10/MWh of electricity.
5. The city has several large hospitals that need a stable and reliable energy source.
6. The city prefers renewable energy. If renewable energies generate more than 30% of the electricity, there will be a renewable tax credit of \$1/MWh for the electricity that is generated by renewable energies.
7. The city prefers energy sources whose cost is stable.
8. The city is concerned with water pollution.

	Energy Plan				Cost	Waste disposal cost	Carbon tax credit	Renewable tax	Total
	Coal	Wind	Nuclear	Hydro					
<b>Plan 1</b>	40%	20%	20%	20%	\$840,000K	\$14,400K	\$48,000K	\$9,600K	\$892,800K
<b>Plan 2</b>	20%	40%	20%	20%	\$912,000K	\$0	\$0	\$11,000K	\$901,000K
<b>Plan 3</b>	20%	20%	40%	20%	\$984,000K	\$14,400K	\$0	\$9,600K	\$988,800K
<b>Plan 4</b>	20%	20%	20%	40%	\$984,000K	\$0	\$0	\$11,000K	\$973,600K

Figure 6.2: Collaboration task description.

represent the values associated with their energy source in contrast to the rest, e.g. coal energy was paired with an economical energy perspective. The team collaborative task was to select a single energy plan and writing a proposal arguing in favor of their decision with respect to the trade-offs, meaning team members needed to negotiate a prioritization among the city requirements with respect to the advantages and disadvantages they were cumulatively aware of. The set of potential energy plans was constructed to reflect different trade-offs among the requirements, with no plan satisfying all of them perfectly. This ambiguity created an opportunity for intensive exchange of perspectives. The collaboration task is shown in Figure 8.4.

## Experimental Procedure

We designed a four-step experiment (Figure 6.3):

- *Step 1. Preparation*

In this step, each worker was asked to provide a nickname, which would be used in the deliberation and collaboration phases. To prepare for the Jigsaw task, each worker was randomly assigned to read an instructional article about the pros and cons of a single energy source. Each article was approximately 500 words, and covered one of four energy sources (coal, wind, nuclear and hydro power). To strengthen their learning and prepare them for the proposal writing, we asked them to complete a quiz reinforcing the content of their assigned article. The quiz was 8 single-choice questions, and feedback including correct answers and explanations was provided along with the quiz.

- *Step 2. Pre-task*

In this step, we asked each worker to write a proposal to recommend one of the four energy sources (coal, wind, nuclear and hydro power) for a city given its five requirements, e.g. “The city prefers a stable energy”. After each worker finished this step, their proposal was automatically posted in a forum as the start of a thread with the title “[Nickname]’s Proposal”.

- *Step 3. Deliberation*

In this step, workers joined a threaded forum discussion akin to those available in many online environments. Each proposal written by the workers in the Pre-task (Step 2) was displayed for workers to read and comment on. Each worker was required to write at least five replies to the proposals posted by the other workers. To encourage the workers to discuss transactively, the task instruction included “when replying to a post, “please elaborate, build upon, question or argue against the ideas presented in that post, drawing from the argumentation in your own proposal where appropriate.”

- *Step 4. Collaboration*

In the collaboration step, team members in a group were first synchronized and then directed to a shared Etherpad<sup>1</sup> to write a proposal together to recommend one of four suggested energy plans based on a city’s eight requirements (Figure 8.4). Etherpad-lite is an open-source collaborative editor [209], meaning workers in the same team were able to see each other’s edits in real-time. They were able to communicate with each other using

<sup>1</sup><http://etherpad.org/>

a synchronous chat utility on the right sidebar. The collaborative task was designed to contain richer information than the individual proposal writing task in Pre-task (Step 2).

Workers were also required to fill out a survey measuring their perceived group outcomes after collaboration.

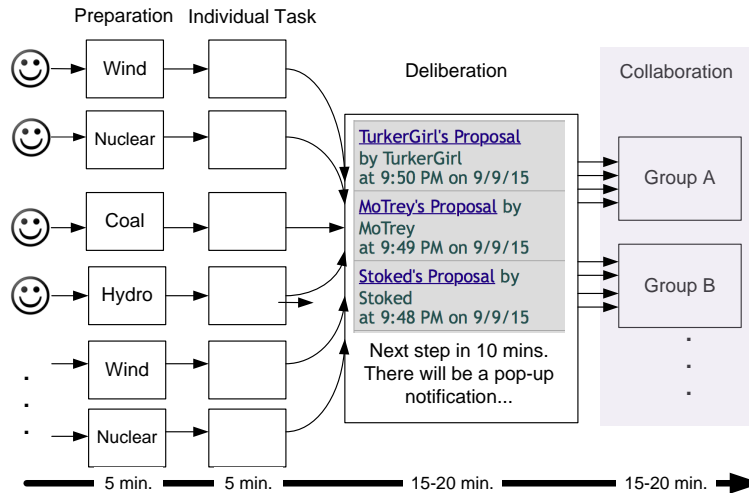


Figure 6.3: Illustration of experimental procedure and worker synchronization for our experiment

### Outcome Measures

Both of our research questions made claims about team success. We evaluated this success using two types of outcomes, namely objective success through quantitative task performance and process measures as well as subjective success through a group satisfaction survey.

The quantitative task performance measure was an evaluation of the quality of the proposal produced by the team. In particular, the scoring rubric (APPENDIX A) defined how to identify the following elements for a proposal:

1. Which requirements were considered
2. Which comparisons or trade-offs were made
3. Which additional valid desiderata were considered beyond stated requirements
4. Which incorrect statements were made about requirements

Positive points were awarded to each proposal for correct requirements considered, comparisons made, and additional valid desiderata. Negative points were awarded for incorrect statements and irrelevant material. We measured **Team Performance** by the total points assigned to the team proposal. Two PhD students who were blind to the conditions applied the rubric to five proposals (a total of 78 sentences) and the inter-rater reliability was good (Kappa = 0.74). The two raters then coded all the proposals.

We used the **length of chat discussion** during teamwork as a measure of team process in the Collaboration step.

**Group Experience Satisfaction** was measured using a four item group experience survey administered to each participant after the Collaboration step. The survey was based on items

used in prior work [41, 70, 115]. In particular, the survey instrument included items related to:

1. Satisfaction with team experience
2. Satisfaction with proposal quality
3. Satisfaction with the communication within the group
4. Perceived learning through the group experience

Each of the items was measured on a 1-7 Likert scale.

#### *Individual Performance Measure*

A score was assigned to each worker to measure their **Individual Performance** during the Collaboration step. The Individual Performance score was based on the contributions they made to their group’s proposal. Similar to team proposal score, a worker received positive points for correct requirements considered, comparisons/trade-offs made, and additional valid desiderata; and negative points for incorrect statements and irrelevant material.

#### *Control Variables*

Intuitively, workers who display more effort in the Pre-task might perform better in the collaboration task. We used the **average group member’s Pre-task proposal length** as a control variable for group performance. We used the worker’s **individual Pre-task proposal length** as a control variable for Individual Performance.

#### *Transactivity Annotation, Prediction and Measurement*

To enable us to use counts of transactive contributions as evidence to inform an automated group assignment procedure, we needed to automatically judge whether a reply post in the Deliberation step was transactive or not using machine learning. Using a validated and reliable coding manual for transactivity from prior work [79], an annotator previously trained to apply that coding manual annotated 426 reply posts collected in pilot studies we conducted in preparation for the studies reported in this chapter. Each of those posts was annotated as either “transactive” or “non-transactive”. 70% of them were transactive.

A transactive contribution displays the author’s reasoning and connects that reasoning to material communicated earlier. Two example posts illustrating the contrast are shown below:

Transactive: “*Nuclear energy, as it is efficient, it is not sustainable. Also, think of the disaster probabilities*”.

Non-transactive: “*I agree that nuclear power would be the best solution*”.

Automatic annotation of transactivity has been reported in the Computer Supported Collaborative Learning literature. For example, researchers have applied machine learning using text, such as chat data [90] and transcripts of whole group discussions [6]. We trained a Logistic Regression model with L2 regularization using a set of features, which included unigrams (i.e., single word features) as well as a feature indicating the post length [64]. We evaluated our classifier with a 10-fold cross validation and achieved an accuracy of 0.843 and a 0.615 Kappa. Given the adequate performance of the model, we used it to predict whether each reply post in the Deliberation step was transactive or not.

To measure **the amount of transactive communication** between two participants in the Deliberation step, we counted the number of times both their posts in a same discussion thread were transactive; or one of them was thread starter and the other participant’s reply was transactive.

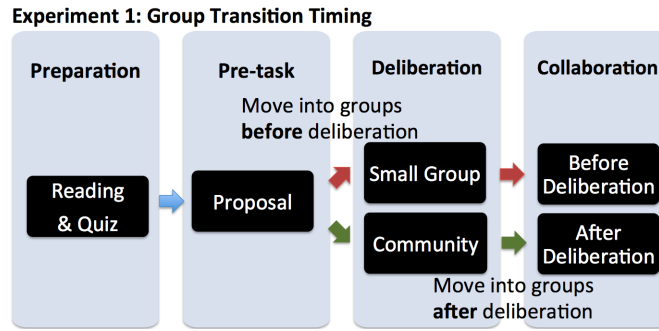


Figure 6.4: Workflow diagrams for Experiment 1.

### 6.3 Experiment 1. Group Transition Timing: Before Deliberation vs. After Deliberation

This experiment (Figure 6.4) assessed whether a measurable improvement would occur if team members transition into groups after community-level deliberation. We manipulated the step in which workers began to work within their small group. To control for timing of synchronization and grouping, in both conditions, workers were synchronized and assigned into small groups based on a Jigsaw paradigm after the Pre-task. The only difference was that for the After Deliberation condition, in the Deliberation step workers could potentially interact with workers both inside and outside their group (40 - 50 workers). Workers were not told that they had been assigned into groups until the Collaboration step (Step 4). 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 only interact with their teammates. By comparing these two conditions, we test our hypothesis that exposure to deliberation within a larger community will improve group performance.

Mechanical Turk does not provide a mechanism to bring several workers to a collaborative task at the same time. We built on earlier investigations that described procedures for assembling multiple crowd workers on online platforms to form synchronous on-demand teams [41, 109]. Our approach was to start the synchronous step at fixed times, announcing them ahead of time in the task description and allowing workers to wait before the synchronous step. A countdown timer in the task window displayed the remaining time until the synchronous step began, and a pop-up window notification was used to alert all participants when the waiting period had elapsed.

#### 6.3.1 Participants

Participants were recruited on MTurk with the qualifications of having a 95% acceptance rate on 1000 tasks or more. Each worker was only allowed to participate once. A total of 252 workers participated in Experiment 1, the workers who were not assigned into groups or did not successfully complete the group satisfaction survey were excluded from our analysis. Worker sessions lasted on average 34.8 minutes. Each worker was paid \$4. To motivate participation



during Collaboration step, workers were awarded a bonus based on their level of interaction with their groups (\$0.1 - \$0.5), while an extra bonus was given to workers whose group submitted a high quality proposal (\$0.5). We included only teams of 4 workers in our analysis, there were in total 22 Before Deliberation groups and 20 After Deliberation groups.

A chi-squared test revealed no significant difference in worker attrition between the two conditions. We considered a worker as having “dropped out” from their team if they were assigned into a team but did not edit the proposal in the Collaboration step. There was no significance difference in the dropout rate of workers between the two conditions ( $\chi^2(1) = 0.08, p = 0.78$ ). The dropout rate for workers in Before Deliberation groups was 30%. The dropout rate for workers in After Deliberation groups was 28%.

### 6.3.2 Results

*Teams exposed to community deliberation prior to group work demonstrate better team performance.*

We built an ANOVA model with Group Transition Timing (Before Deliberation, After Deliberation) as the independent variable and Team Performance as the dependent variable. In order to control for differences in average verbosity across teams, we included as a covariate for each group the Pre-task proposal length averaged across team members. There was a significant main effect of Group Transition Timing on Team Performance ( $F(1,40) = 5.16, p < 0.05$ ) such that After Deliberation groups had a significantly better performance ( $M = 9.25, SD = 0.87$ ) than the Before Deliberation groups ( $M = 6.68, SD = 0.83$ ), with an effect size of 2.95 standard deviations.

We also tested whether the differences in teamwork process between conditions was visible in the extent of the length of chat discussion. We built an ANOVA model with Group Transition Timing (Before Deliberation, After Deliberation) as the independent variable. In this case, there was no significant effect. Thus, teams in the After Deliberation condition were able to achieve better performance in their team product without requiring more discussion.

*Workers exposed to community deliberation prior to group work demonstrate higher-quality individual contributions.*

In addition to evaluating the overall quality of the team performance, we also assigned an Individual Performance score to each worker based on their own contribution to the team proposal. There was a significant correlation between the Individual Performance and the corresponding Team Performance ( $r = 0.35, p < 0.001$ ), suggesting that the improvements in After Deliberation team product quality were at least in part explained by benefits individual workers gained through their exposure to the community during the deliberation phase. To assess that benefit directly, we built an ANOVA model with Group Transition Timing (Before Deliberation, After Deliberation) as the independent variable and Individual Performance as the dependent variable. TeamID and assigned energy condition (Coal, Wind, Hydro, Nuclear) were included as control variables nested within condition. Additionally, Individual Pre-task proposal length was included as a covariate. There was no significant main effect of Group Transition Timing ( $F(1,86) = 2.4, p = 0.12$ ) on Individual Performance, although there was a trend such that workers in the Before Deliberation condition ( $M = 1.05, SD = 0.21$ ) contributed less than the workers in the After Deliberation condition ( $M = 1.75, SD = 0.23$ ). There was only a marginal effect of individual Pre-task proposal length ( $F = 3.41, p = 0.06$ ) on Individual Performance, again suggesting that

the advantage to team product quality was at least in part explained by the benefits individual workers gained by their exposure to the community during the deliberation, though the strong and clear effect appears to be at the team level. There was no significant effect of assigned energy.

#### *Survey results*

In addition to assessing group and individual performance using our scoring rubric, we assessed the subjective experience of workers using the group experience survey discussed earlier. For each of the four aspects of the survey, we built an ANOVA model with Group Transition Timing (Before Deliberation, After Deliberation) as the independent variable and the survey outcome as the dependent variable. TeamID and assigned energy condition (Coal, Wind, Hydro, Nuclear) were included as control variables nested within condition. There were no significant effects on any of the subjective measures in this experiment.

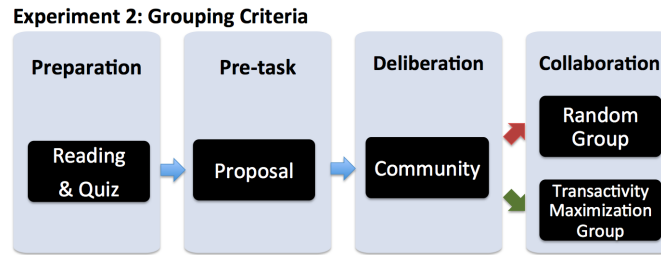


Figure 6.5: Workflow diagrams for Experiment 2.

## 6.4 Experiment 2. Grouping Criteria: Random vs. Transactivity Maximization

While in Experiment 1 we investigated the impact of exposure to community resources prior to teamwork on team performance, in Experiment 2 we investigated how the nature of the experience in that context may inform effective team formation, leading to further benefits from that community experience on team performance. This time teams in both conditions were grouped after experiencing the deliberation step in the community context. In Experiment 1, workers were randomly grouped based on the Jigsaw paradigm. In Experiment 2 (Figure 6.5), we again made use of the Jigsaw paradigm, but in the experimental condition, which we termed the Transactivity Maximization condition, we additionally applied a constraint that preferred to maximize the extent to which workers assigned to the same team had participated in transactive exchanges in the deliberation. In the control condition, which we termed the Random condition, teams were formed by random assignment. In this way we tested the hypothesis that observed transactivity is an indicator of potential for effective team collaboration.

From a technical perspective in Experiment 2 we manipulated how the teams were assigned. For the Transactivity Maximization condition, teams were formed so that the amount of transactive discussion among the team members was maximized. A Minimal Cost Max Network Flow algorithm was used to perform this constraint satisfaction process [5]. This standard network flow algorithm tackles resource allocation problems with constraints. In our case, the constraint

was that each group should contain four people who have read about different energies (i.e. a Jigsaw group). At the same time, the minimal cost part of the algorithm maximized the transactive communication that was observed among the group members during Deliberation step. The algorithm finds an approximately optimal grouping within  $O(N^3)$  ( $N$  = number of workers) time complexity. A brute force search algorithm, which has an  $O(N!)$  time complexity, would take too long to finish since the algorithm needs to operate in real time. Except for the grouping algorithm, all the steps and instructions were identical for the two conditions.

### 6.4.1 Participants

A total of 246 workers participated in Experiment 2, the workers who were not assigned into groups or did not complete the group satisfaction survey were excluded from our analysis. Worker sessions lasted on average 35.9 minutes. We included only teams of 4 workers in our analysis. There were in total 27 Transactive Maximization teams and 27 Random teams, with no significant difference in attrition between conditions ( $\chi^2(1) = 1.46, p = 0.23$ ). The dropout rate of workers in Random groups was 27%. The dropout rate of workers in Transactivity Maximization groups was 19%.

### 6.4.2 Results

As a manipulation check, we compared the average amount of transactivity observed among teammates during the deliberation between the two conditions using a t-test. The groups in the Transactive Maximization condition ( $M = 12.85, SD = 1.34$ ) were observed to have had significantly more transactive communication during the deliberation than those in the Random condition ( $M = 7.00, SD = 1.52$ ) ( $p < 0.01$ ), with an effect size of 3.85 standard deviations, demonstrating that the maximization was successful in manipulating the average experienced transactive exchange within teams between conditions.

*Teams that experienced more transactive communication during deliberation demonstrate better team performance.*

To assess whether the Transactivity Maximization condition resulted in more effective teams, we tested for a difference between group formation conditions on Team Performance. We built an ANOVA model with Grouping Criteria (Random, Transactivity Maximization) as the independent variable and Team Performance as the dependent variable. Average team member Pre-task proposal length was again the covariate. There was a significant main effect of Grouping Criteria ( $F(1,52) = 6.13, p < 0.05$ ) on Team Performance such that Transactivity Maximization teams ( $M = 11.74, SD = 0.67$ ) demonstrated significantly better performance than the Random groups ( $M = 9.37, SD = 0.67$ ) ( $p < 0.05$ ), with an effect size of 3.54 standard deviations, which is a large effect.

Across the two conditions, observed transactive communication during deliberation was significantly correlated with Team Performance ( $r = 0.26, p < 0.05$ ). This also indicated teams that experienced more transactive communication during deliberation demonstrated better team performance.

*Teams that experienced more transactive communication during deliberation demonstrate more intensive interaction within their teams.*

In Experiment 2, workers were assigned to teams based on observed transactive communication during the deliberation step. Assuming that individuals that were able to engage in positive collaborative behaviors together during the deliberation would continue to do so once in their teams, we would expect to see evidence of this reflected in their observed team process, whereas we did not see such an effect in Experiment 1 where teams were assigned randomly in all conditions. Group processes have been demonstrated to be strongly related to group outcomes in face-to-face problem solving settings [199]. Thus, we should consider evidence of a positive effect on group processes as an additional positive outcome of the experimental manipulation.

In order to test whether such an effect occurred in Experiment 2, we built an ANOVA model with Grouping Criteria (Random, Transactivity Maximization) as the independent variable and length of chat discussion during teamwork as the dependent variable. There was a significant effect of Grouping Criteria on length of discussion ( $F(1,45) = 9.26, p < 0.005$ ). Random groups ( $M = 20.00, SD = 3.58$ ) demonstrated significantly shorter discussions than Transactive Maximization groups ( $M = 34.52, SD = 3.16$ ), with an effect size of 4.06 standard deviations.

Table 6.1 shows one transactive and one non-transactive collaboration discussion. The transactive discussion contains reasoning about the pros and cons of the energy plans, which can easily translate into the team proposal. The non-transactive collaborative discussion try to come to a quick consensus without discussing each participant's rationale of choosing an energy plan. Then to write a team proposal, the participants need to come up and organize their reasons of choosing the plan. For participants who initially did not pick the chosen energy plan, without the transactive reasoning process, it is difficult for them to integrate their knowledge and perspective into the team proposal.

A transactive discussion example	A non-transactive discussion example
<p>A: based on plan 1 and 2 I am thinking 2 only because it reduces greenhouse gases</p> <p>B: Yeah so if we go with 2, we will need to trade off the water pollution and greenhouse gas</p> <p>C: BUT we run into the issue of budget... so where do we say the extra almost \$100k comes from?</p>	<p>A: My two picks are Plan 1 and Plan 2</p> <p>B: Alright, lets take a vote. Type either Plan 1 or Plan 2 in chat.</p> <p>B: Plan 2</p> <p>C: Plan 1</p> <p>D: plan 2.</p> <p>B: That settles it, its plan 2.</p>

Table 6.1: Transactive vs. Non-transactive Discussions during Team Collaboration.

*Workers whose groups formed based on transactive communication during the deliberation process demonstrate better individual contributions to the team product.*

In Experiment 2, again we were interested in the contribution of individual performance effect to group performance effect. There was a significant correlation between Individual Performance and the corresponding Team Performance ( $r = 0.34, p < 0.001$ ), suggesting again that the advantage to team product quality was at least in part explained by the synergistic benefit individual workers gained by working with team members they had previously engaged during the community deliberation.

To assess that benefit directly, we built an ANOVA model with Grouping Criteria (Random, Transactivity Maximization) as the independent variable and Individual Performance as

the dependent variable. TeamID and assigned energy condition (Coal, Wind, Hydro, Nuclear) were included as control variables nested within condition. Additionally, individual Pre-task proposal length was included as a covariate. There was a marginal main effect of Grouping Criteria ( $F(1,166) = 3.70, p = 0.06$ ) on Individual Performance, such that workers in the Random condition ( $M = 1.64, SD = 0.21$ ) contributed less than the workers in the Transactivity Maximization condition ( $M = 2.32, SD = 0.20$ ), with effect size 3.24 standard deviations. There was a significant effect of individual Pre-task proposal length ( $F = 7.86, p = 0.0005$ ) on Individual Performance, again suggesting that the advantage to team product quality was at least in part explained by the benefit individual workers gained by working together with the peers they had engaged during the deliberation phase, though the strong and clear effect appears to be at the team performance level. There was no significant effect of assigned energy.

#### *Survey results*

For each of the four aspects of the group experience survey, we built an ANOVA model with Grouping Criteria (Random, Transactivity Maximization) as the independent variable and the survey outcome as the dependent variable. TeamID and assigned energy condition (Coal, Wind, Hydro, Nuclear) were included as control variables nested within condition. There were no significant effects on Satisfaction with team experience or with proposal quality. However, there was a significant effect of condition on Satisfaction with communication within the group ( $F(1,112) = 4.83, p < 0.05$ ), such that workers in the Random teams ( $M = 5.12, SD = 1.7$ ) rated the communication significantly lower than those in the Transactivity Maximization teams ( $M = 5.69, SD = 1.51$ ), with effect size .38 standard deviations. Additionally, there was a marginal effect of condition on Perceived learning ( $F(1,112) = 2.72, p = 0.1$ ), such that workers in the Random teams ( $M = 5.25, SD = 1.42$ ) rated the perceived benefit to their understanding they received from the group work lower than workers in the Transactivity Maximization teams ( $M = 5.55, SD = 1.27$ ), with effect size 0.21 standard deviations. Thus, with respect to subjective experience, we see advantages for the Transactivity Maximization condition in terms of satisfaction of the team communication and perceived learning, but the results are weaker than those observed for the objective measures. Nevertheless, these results are consistent with prior work where objectively measured learning benefits are observed in high transactivity teams [51].

## **6.5 Implications for team-based MOOCs**

One potential context for the application of the results of this work is that of team based MOOCs, such as those offered on the NovoEd platform<sup>2</sup>, and soon on edX<sup>3</sup> through the new team platform extensions. Since introduction in 2011, MOOCs have attracted millions of learners. However, social isolation is still the norm for current generation MOOCs. Team-based learning can be a beneficial component in online learning [123]; how to do that in the context of MOOCs is an open question. Two major constraints, in the context of MOOCs, are that it can be technically very difficult to mediate and support large-scale conversations, and the global learner population from various time zones is a challenge to synchronous communication [111]. Group projects require that learners be present on a particular schedule, reducing the flexibility and convenience factor

<sup>2</sup>A team based MOOC platform, <https://novoed.com/>

<sup>3</sup><https://www.edx.org/>

in online study and possibly causing anxiety and/or resentment, particularly if the purpose of the group work is not clear and the group experience is not positive. It is important to consider, however, that it may be more difficult to coordinate students in MOOCs than in crowd work because of the desire to align every student with a group, as opposed to grouping whichever subset of workers happens to be available at a given time. Instructors will most likely have to encourage students to arrive within pre-stated time periods and it may be necessary to modify MOOC pages to alert students to upcoming group activities. One method of ensuring learner participation in online collaboration is to demonstrate the value of group learning by assessing (defined here as assignment of a grade) both the product and process of group work [167].

## 6.6 Discussions

There are three underlying mechanisms that make the transactivity maximization team formation work. The first is that the evidence that participants can transactively discuss with each other shows that they respect each other's perspective. This process indicates they can effectively collaborate with each other in a knowledge integration task. The second mechanism is the effect that the students have read each other's individual work. In our experiments, the individual task is similar to the collaboration task. This helps the two participants understand each other's knowledge and perspective. Thirdly, the community deliberation process is also a self-selection process. Participants tend to comment on posts they find more interesting. This may also contribute to their future collaboration.

## 6.7 Conclusions

In this chapter we present two experiments to address two related research questions. The first question was whether participation in deliberation within a community is more valuable as preparation for teamwork than participation in deliberation within the team itself. Here we found that deliberation within a community has advantages in terms of the quality of the product produced by the teams. There is evidence that individuals may also contribute higher quality individual contributions to their teams as a further result, though this effect is only suggestive. We see no effect on subjective survey measures.

The second related question was the extent to which additional benefit from participation in the deliberation in a community context to teams could be achieved by using evidence of potential successful team interactions from observed transactive exchanges during the deliberation. Here we found that teams formed such that observed transactive interactions between team members in the deliberation was maximized produced objectively higher quality team products than teams assigned randomly. There was also suggestive evidence of a positive effect of transactivity maximization on individual contribution quality. On subjective measures we see a significant positive impact of transactivity maximization on perceived communication quality and a marginal impact perceived enhanced understanding, both of which are consistent with what we would expect from the literature on transactivity where high transactivity teams have been demonstrated to produce higher quality outcomes and greater learning [90, 172].

These results provide positive evidence in favor of a design for a team formation strategy in two stages: Individuals first participate in a pre-teamwork deliberation activity where they explore the space of issues in a context that provides beneficial exposure to a wide range of perspectives. Individuals are then grouped automatically through a transactivity detection and maximization procedure that uses communication patterns arising naturally from community processes to inform group formation with an aim for successful collaboration.





## Chapter 7

# Team Collaboration Communication Support: Study 5

The corpus analysis in Study 2 indicates we can either support the collaboration communication or improve group formation to enhance team collaboration product quality. In Study 3 and 4, we first explored supporting online team formation. In Study 5, we will study how to support online synchronous collaboration communication with a conversational agent. There are two hypothesis in Study 5. First, we test the hypothesis where a transactive based team formation can improve students' learning gain:

*H1. Students in teams that are formed based on evidence of transactive discussions during deliberation have higher learning gain.*

Second, to understand whether communication collaboration support, i.e. discussion facilitation agent, can be helpful in online collaborative context. In this chapter, we focus on supporting synchronous discussion through a revised version of the conversation agent Bazaar [1]. We test the hypothesis that:

*H2. Groups that are supported by communication facilitation agent during deliberation have higher activity level and better reasoning discussion during collaboration, and therefore students have higher learning gain.*

Figure 7.1 presents our hypotheses in this chapter.

In Study 5, we use the same experimental workflow and knowledge integration task as in Study 3 and 4. Each participant is exposed to reading materials of one type of energy and asked to propose an energy plan for a city based on his/her perspective. After the group deliberation process, in which participants were asked to comment on each others proposal, they were then assigned to groups to complete a collaborative task, in which they need to come up with a group proposal while at the same time discuss in Bazaar to discuss trade-offs and reach a consensus. In order to maximize team collaboration product quality, ideally group members are expected to articulate the trade-off between plans from their perspective in the discussion. So that the group is meant to work together to develop a more extensive plan taking all factors into consideration. The purpose of the conversation agent is to intensify the interaction between participants, and encourage them to elaborate on their argument from their own perspective.

In this study, we hope to observe the effect of conversational agent support on group performance and individual learning, and further investigate whether theres a difference in the effect of

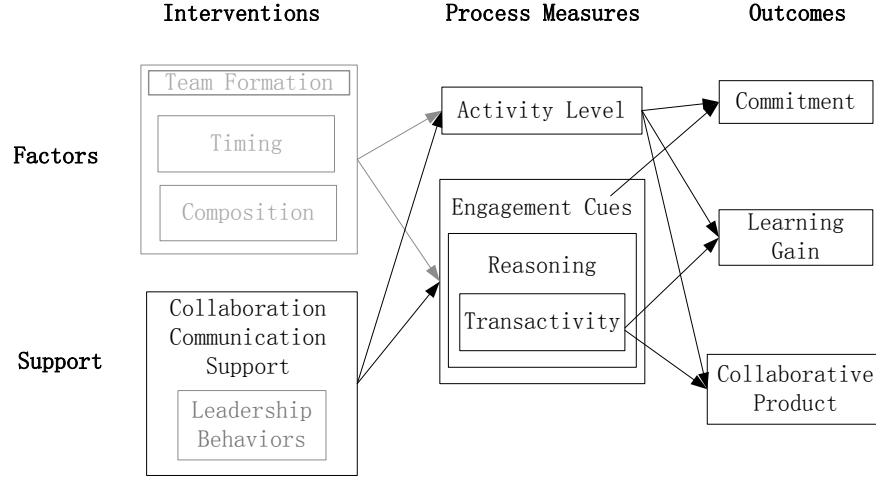


Figure 7.1: Chapter hypothesis.

the support between high transactivity groups and low transactivity groups. We adopted a 2x2 experimental design. The first experimental variation is whether the students are assigned to groups based on the transactivity optimization algorithm or that they are assigned to groups randomly, which is an indicator of high or low transactivity groups. The second experimental variation is whether students will receive scripted support from a conversational agent in their group collaboration discussion or not. We consider our this MTurk experiment with conversational agent support as a starting point to support synchronous discussion in team-based MOOCs.

In addition to participants collective performance on the group task, we are testing whether the experiment setting would cause a difference in students individual domain-specific knowledge acquisition, i.e. learning gain. In order to capture students individual learning of domain-specific knowledge, we administered a pre-test and post-test. The measure of learning in this task will be individual learning gain from pre-test and post-test.

## 7.1 Bazaar Framework

Bazaar is a conversational agent that can automatically facilitate synchronous collaborative learning [1]. The publicly available Bazaar architecture enables easy integration of a wide variety of discussion facilitation behaviors. During team collaboration, the agent takes the role of a teacher and make facilitation moves. Introduction of such technology in a classroom setting has consistently led to significant improvements in student learning [1], and even positive impacts on the classroom environment outside of the collaborative activities [39].

In order to provide a very light form of micro-level script-based support for the collaboration, the computer agent facilitator in our design kept track of the students in the room and prompts that had been given to the students. For each student, it tracked the prompts that have been provided to him. It also tracked which plans the student has mentioned. This tracking was done to make sure each student could engage with various prompts of the reflection exercise and no student saw the same question prompt twice.

## 7.2 Prompts

The prompts aim to facilitate group members to elaborate on their ideas and intensify the interaction between different group members from different perspectives. We consider that by providing conversational support in synchronous discussion, we will observe better group processes and better knowledge acquisition for individual students. Bazaar analyse the event stream in search of triggers for supportive interventions. After each student entered one chat line, Bazaar will automatically identify whether the turn contains reasoning or not.

In the preparation step, each participant is assigned an energy and a perspective that is associate with that energy:

Coal: energy that is most economical

Wind: most environmentally friendly and with lowest startup costs

Nuclear: economically in the long run

Hydro: environmental friendly and reliable

When the person doesnt show reasoning, there are two kinds of prompts: 1) Ask the person to elaborate on the plan from his perspective that was assigned in the preparation step, there are three templates: a. Hey (NAME), can you elaborate on the reason you chose plan (N) from your perspective of what would be (NAME's PERSPECTIVE)?

b. Hey (NAME), can you be more specific about why you chose plan (N) from your perspective of (NAME's PERSPECTIVE)?

c. Hey (NAME), what do you think are the pros and cons of plan (N) from your perspective of (NAME's PERSPECTIVE)?

2) Ask the person to compare the plan with another plan that hasnt been fully discussed (i.e., elaborate the pros and cons).

a. Hey (NAME1), you have proposed plan (N), and (NAME2) has proposed plan (N). What do you think are the most important trade-offs between the two plans in terms of your perspective of (NAME's PERSPECTIVE)?

When the person shows reasoning, there are two kinds of prompts: 1) Ask someone else who has proposed a different plan which hasnt been fully discussed to compare the two plans.

a. Hey (NAME), you have proposed plan (N), and (NAME) has proposed plan (N). Can you compare the two plans from your perspective of (NAME's PERSPECTIVE)?

2) Ask someone else to evaluate this plan, there are three templates:

a. Hey (NAME1), can you evaluate (NAME2)s plan from your perspective of (NAME's PERSPECTIVE)?

b. Hey (NAME1), how do you like (NAME2)s plan from your perspective of NAME's PERSPECTIVE?

c. (NAME1), would you recommend (NAME2)s plan from your perspective of NAME's PERSPECTIVE?

When theres nobody talking in the chat, it will ask someone who hasnt been prompted, there are two possible prompt templates, 1) Hey (NAME), which of the plans seems to be the best from your perspective of (NAME's PERSPECTIVE)?

2) Hey (NAME), which plan do you recommend from your perspective of (NAME's PERSPEC-

TIVE)?

During a synchronous chat collaborative activity, the students may not be motivated to start discussing and may engage in goalless interactions or no interactions at all. Bazaar also offers starter prompt and reminder prompt.

Starter prompt:

Use this space to discuss the relative merits and demerits of each plan from your assigned perspective in step 2. That was the step where you were asked to write an individual proposal and post it to the discussion forum. Work towards a consensus on which plan you think is the best as a group. Wherever possible, supplement plan suggestions, agreements and disagreements with reasons, and try to refer to or build upon your teammates' ideas. I will support you in the process.

Reminder prompt when theres five minutes left:

Thank you for participating in the discussion. You now have just 5 minutes to finalize a joint argument in favor of one plan and complete your proposal on the left. Keep in mind that you will be evaluated solely based on the quality of your proposal (i.e., the thoroughness of the reasoning displayed).

Here are several examples of the prompts provided by Bazaar. If someone proposes a plan without providing reasoning, Bazaar will prompt him with "Hey Shannie, can you elaborate on the reason you chose plan 2 from your perspective of what would be most economical?" On the other hand, if a group member proposes a plan with reasoning, and a second group member proposes a different plan, Bazaar will prompt the second team member to compare the two plans, e.g., "Hey Beth, you have proposed plan 1, and Corey has proposed plan 3, can you compare the two plans from your perspective of environmental friendliness and reliability?"

## 7.3 Learning Gain

To test the hypothesis where a transactive based team formation and the team collaboration support can improve students' learning gains. We added one pre-test step and one post-test step to the crowdsourcing experiment workflow. The task in the pre and post test is similar to the collaborative task, which is to design an energy plan for a city based on its requirements. The task description of the pre and post-test is the same, which is listed below.

The proposals in the pre and post-test were scored according to the same coding manual. Participant's learning gain is measured as the difference from the pre-test to post-test.

## 7.4 Participants

Participants were recruited on MTurk with the qualifications of having a 95% acceptance rate on 1000 tasks or more. Each worker was only allowed to participate once. We included only teams of 4 workers in our analysis. A total of 67 four-people teams were formed (Table 7.1). Each Turker was paid \$8. To motivate participation during Collaboration step, workers were awarded a bonus based on their level of interaction with their groups (\$0.1 - \$0.5), while an extra bonus was given to workers whose group submitted a high quality proposal (\$0.5).

Complete the following task with your best knowledge:

Read the following instructions carefully, and write a proposal based on your best knowledge. You are not allowed to refer to external resources in this task.

Diamond City is a city undergoing rapid economic growth, and thus is in need of large volume of reliable energy. It is a busy city with a high volume of traffic and a growing population. It is home to a famous tech-oriented university, which means the city has access to a sophisticated workforce. The city is in a windy area, and is rich in natural resources, including fossil fuels and water. However, based on the city's commitment to development, it has a limited budget for financing energy.

Please write a 80-120-word energy proposal for Diamond City. You need to choose from four types of energy: A. coal energy B. wind energy C. nuclear energy D. hydroelectric energy. You can either choose one or any combination of the four. Please be specific about your reasons. Explain your arguments well, and demonstrate awareness of differing priorities and the pros and cons of the different sources in relation to the characteristics of the city. Your proposal will be evaluated on its thoughtfulness and comprehensiveness. You will get \$0.5 bonus if your proposal achieves a top ranking score.

Figure 7.2: Instructions for Pre and Post-test Task.

	Team Formation		
		Random	Transactivity Maximization
<b>Communication Support</b>	Without Support	16	16
	With Support	16	16

Table 7.1: Number of teams in each experimental condition.

## 7.5 Results

### 7.5.1 Team Process and Team Performance

Controlling for the average proposal length during individual work, teams in the transactivity maximization condition had significant better performance than the teams that were formed randomly ( $F(1,55) = 5.39$ ,  $p = 0.02$ ).

We observed a marginal effect of collaboration support, and a significant interaction between collaboration discussion support and team formation method. Arguably the teams that were formed randomly need the most support for their collaboration discussion. Then the randomly formed teams would benefit more from Bazaar than the Transactivity maximization teams. Instead, the results indicate the opposite.

For comparisons/trade-offs made in the proposal, there's a marginal effect of team formation and a significant effect of collaboration support and a significant interaction the combination of the two is significantly better than everything else.

Controlling for the average proposal length during individual work, teams in the Transactivity Maximization Condition talk more ( $p = 0.02$ ) and contribute more transactivity ( $F(1,27) = 5.39$ ,  $p = 0.03$ ).

Teams with Bazaar support has a marginal higher percentage of reasoning in their discussion

( $F(1,29) = 3.06$ ,  $p = 0.09$ ).

### **7.5.2 Learning Gains**

Bazaar has a significant effect on correct tradeoff points in the post test (controlled for correct tradeoff points in pre test). ( $p = 0.041$ )

And the group proposal score is significant in predicting individual learning ( $p = 0.009$ ), which means students who had better team performance also learnt more. And bazaar\* trans groups had the most learning. (though not significant)

## **7.6 Discussions**

### **7.6.1 High Transactivity Groups Benefit More from Bazaar Communication Support**

### **7.6.2 Learning Gain**

## **7.7 Conclusions**

The research reported in this chapter provides evidence of a result when the variables of interest are isolated in a setting in which it is possible to achieve high internal validity. In order for these results to achieve practical value, in the next two chapters we will test the team formation method in a team-based MOOC.

## Chapter 8

# Team Formation Intervention Study in a MOOC: Study 6

The preliminary results from the crowdsourced experiments in Study 3 and 4 were in accordance with our hypothesis, which demonstrated the internal validity of our team formation hypothesis. In order to study how our team formation paradigm will work in a real MOOC, we did two deployment studies, Study 6 and 7. We would like to know (1) whether this team formation process will work in real MOOC environment (2) if we can see evidence of the benefit of our team formation method. Crowd workers likely have different motivations from MOOC students. Only a deployment study can answer questions like: how many MOOC students will actually participate in virtual team collaboration? Whether the small team collaboration in MOOC is enjoyable or not? An MTurk task is usually short, typically lasts less than an hour. However, in an actual MOOC, a virtual team collaboration can be several weeks long. A deployment study will help us identify what other support is needed to support long-term MOOC team collaboration. In this chapter, we present our first team formation intervention study in a real MOOC, where we added an optional Team Track to the original MOOC.

### 8.1 The Smithsonian Superhero MOOC

The MOOC we collaborated with is called “The Rise of Superheroes and Their Impact On Pop Culture”, offered by the Smithsonian Institution. The course is six weeks long and offered on the edX platform. The first offering of the course attracted more than 40,000 enrollment. Students’ learning goal in this MOOC is to get a sense of how comic book creators build a character and draw inspiration both from the world around them and inside them to construct a story. The teaching goal of the course is to understand how the genre of superheroes can be used to reflect the hopes and fears that are salient during periods of American history. In this MOOC, students either design a superhero of their own or write a biography of an existing superhero. On the one hand, we chose this course because it is a social science course, where it is easier to design the deliberation and collaboration tasks that are similar to the ones we have used in the crowdsourced environment. The original course project was also a design task where there was no right or wrong answer. Based on that task, we designed a collaborative design task which contains the

same key components. Similar to the collaborative energy proposal design task in Study 3-5, in the design task, it would be important to respect each other's different perspectives and combine all team members' designs. On the other hand, in the previous run of this MOOC, many students expressed that they need more feedback on their superhero design in post-course survey. We think a small team collaboration would be an opportunity for team members to provide feedback to each other.

Starting from the second offering of the course, there were two tracks of study: Creative track and Historical track. The students in the **Creative Track** submit a "superhero sketchpad" to finish the MOOC. The sketchpad contains five elements:

- Create your superhero's superpowers and weaknesses.
- Detailing your superhero biography by telling an origin story for how your superhero came to be.
- Design a supervillain and your rationale for choosing him/her.
- Construct a story that includes three components: build up, conflict and resolution.
- Build three original comic panels that bring your story to life.

The history-focused track (**Historical Track**) students write an essay based and involves analysis of information and comics. Students in this track write an analysis of a superhero of their choice.

Since it was the first time to run this MOOC with a team component, the instructors were hesitate to drastically change the course. We decided to add an optional Team Track to the two tracks of study: Creative Track and Historical Track.

## 8.2 Data Collection

In Study 6, we will collect students' course community discussion data, team project submissions, teams' interaction trace in team space and post-course survey data. To understand how students are collaborating in these teams, we will qualitatively analyze students' discussion in the team space.

We designed a team project which was to design a superhero team collaboratively with the superhero they designed or chose. To enable us meaningfully compare the team track projects and the individual tracks, the Team Track project also contained the same components. Each student team will submit a Google Doc together as their final team project. In the project, they will discuss three questions among their team:

1. What archetype(s) does your superhero fit into?
2. Will your superhero team support government control or personal freedom?
3. What current social issue will your superhero team take on?

To see what benefits students can gain from joining a team collaboration, we will compare the core components in student project submissions from the individual track and the team track.



## **8.3 Adapting our Team Formation Paradigm to the MOOC**

To prepare for the deployment study, we worked with the course staff and designer together to design the Team Track for the third offering of the superhero MOOC. We customized our team formation paradigm around the need of the MOOC instructors and students. During the actual MOOC, I worked as a course staff. I was responsible for grouping students, sending out team assignment emails, replying to emails and course forum posts that are related to team formation and team collaboration. I served as the “data czar” for SmithsonianX, so I receive course data from edX (including enrollment information and discussion posts) weekly.

Similar to the workflow in our crowdsourced experiments, students in the Team Track went through three steps to finish the course: Individual Work, Course community deliberation, Team Formation and Collaboration.

### **8.3.1 Step 1. Individual Work**

To ensure that students have enough time to do individual work. Students will decide whether to join the team track or not during Week 3 of the course. In Week 1 and 2, students in the team track will work on the same task of designing superhero origin story and superpower as students in the individual track.

### **8.3.2 Step 2. Course community deliberation**

Similar to the deliberation task in our crowdsourced study, students who want to join the Team Track are required to post their individual work to the discussion board and provide feedback to each other. In a real MOOC, we cannot assume that everyone will participate in the team track, we use posting their work to the course discussion board as a sign-up process for the team track. Since providing feedback is naturally transactive, we designed a discussion instruction to encourage students to give each other feedback on their individual work. The instructions are shown in Figure 8.2. A screenshot of students discussion prompt and threads in course week 4 are shown in Figure 8.1.

### **8.3.3 Step 3. Team Formation and Collaboration**

We planned to use the same transactive maximization algorithm to form teams based on their forum discussion. Similar to the different energy conditions in our crowdsourced studies, we also planned to use the original two tracks of study as our Jigsaw conditions. However, since there were not enough students who participated in the Team Track, we ended up form teams based on when the students post in the forum. Students in the Team Track formed teams of at least two students.

We assigned a team space to each team. Since students might be hesitate to post in an empty team space, we posted an initial post in each team space. An example is shown in Figure 8.3. In this post, we encouraged students in the team to discuss about their project in this team space. We also provide each team a url link to a synchronous chat space that was dedicated to each team.

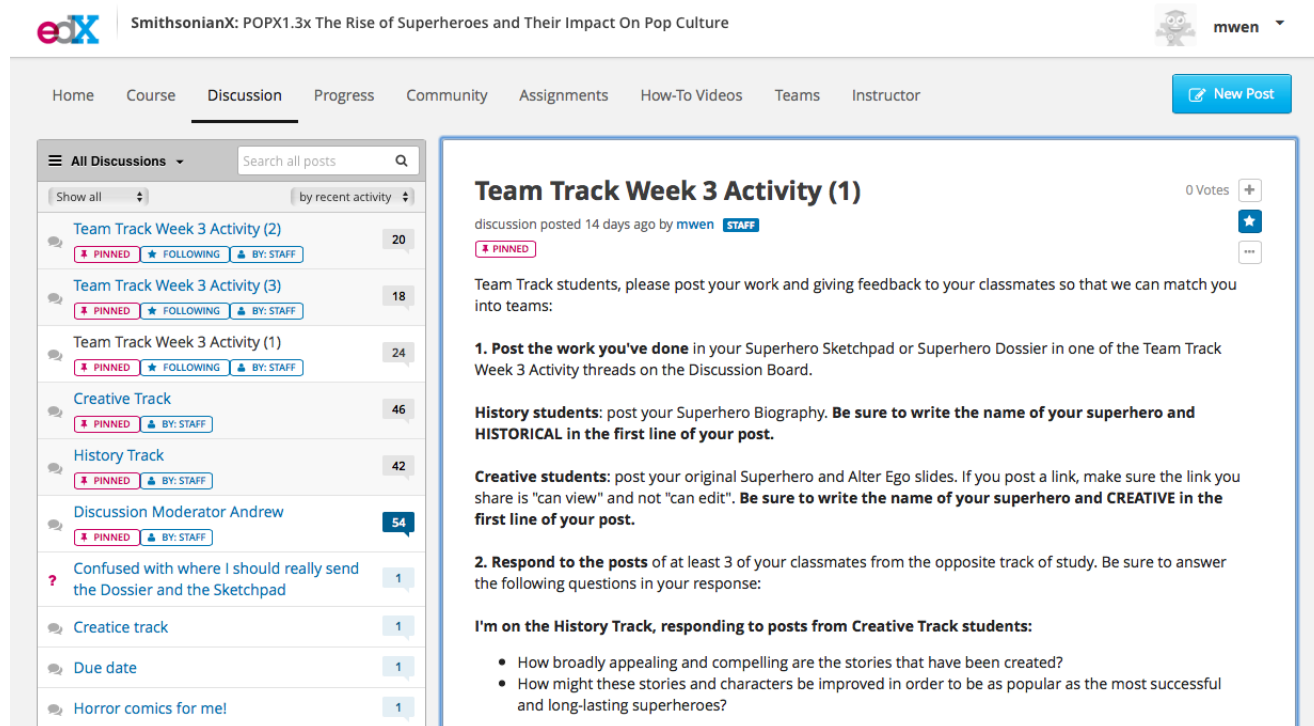


Figure 8.1: A screenshot of Week 4 discussion threads for team formation.

## 8.4 Results

### 8.4.1 Course completion

In total, 6,485 students enrolled in the MOOC, 145 of them were verified. After the initial three weeks less than 940 students were active in a given week. At the end of week 4, there were 16 students who posted and commented in the team formation threads (94 posts in 3 threads). Only one of these students are from the Historical Track. The discussion forum snapshot is shown in Figure 8.1. In total we formed 6 teams, four 3-people teams and two 2-people teams. 4 of the teams submitted their final team projects (Table 9.2).

	Creative + Historical	Team Track
Active in Week 3	1195 (99%)	16 (1%)
Submitted Final Project	180 (94%)	11 (6%)
Awarded Certificates	82 (88%)	11 (12%)

Table 8.1: Student course completion across tracks.

This week, we will put you into teams of 3-4 students. These will be your teams for the remainder of the course. As a team, you will work together in week 4 to discuss your individual work and in weeks 5 and 6 to complete a team final project in which you will create a superhero team.

In order to form the teams, we need to ask you to each post the work you've done in your Superhero Sketchpad or Superhero Dossier to one of the Team Track Week 3 Activity pinned threads in the Discussion Board.

After everyone has responded to at least 3 postings, we will place each student who posted into a team. You will receive an email from us with a link to your team discussion space here on edX during Week 4.

### **Instructions for posting in the Team Track Week 3 Activity thread**

Here are the instructions for posting your work and giving feedback to your classmates so that we can match you into teams:

1. Post the work you've done in your Superhero Sketchpad or Superhero Dossier in one of the Team Track Week 3 Activity threads on the Discussion Board. Be sure to mark whether your work is HISTORICAL or CREATIVE in the first line of your post. e.g. Superman Dossier: Historical Track History students: post your Superhero Biography. Be sure to write the name of your superhero and HISTORICAL in the first line of your post.

Creative students: post your original Superhero and Alter Ego slides. If you post a link, make sure the link you share is "can view" and not "can edit". Be sure to write the name of your superhero and CREATIVE in the first line of your post.

2. Respond to the posts of at least 3 of your classmates from the opposite track of study. Be sure to answer the following questions in your response:

I'm on the History Track, responding to posts from Creative Track students:

How broadly appealing and compelling are the stories that have been created? How might these stories and characters be improved in order to be as popular as the most successful and long-lasting superheroes?

I'm on the Creative Track, responding to posts from History Track students:

How well connected are the values of the chosen superhero and the underlying social issue(s) discussed in the selected news article? How could the chosen superhero be further developed in order to connect better with contemporary social issues?

3. Check your own post to see what feedback you've received.

Figure 8.2: Instructions for Course Community Discussion and Team Formation.

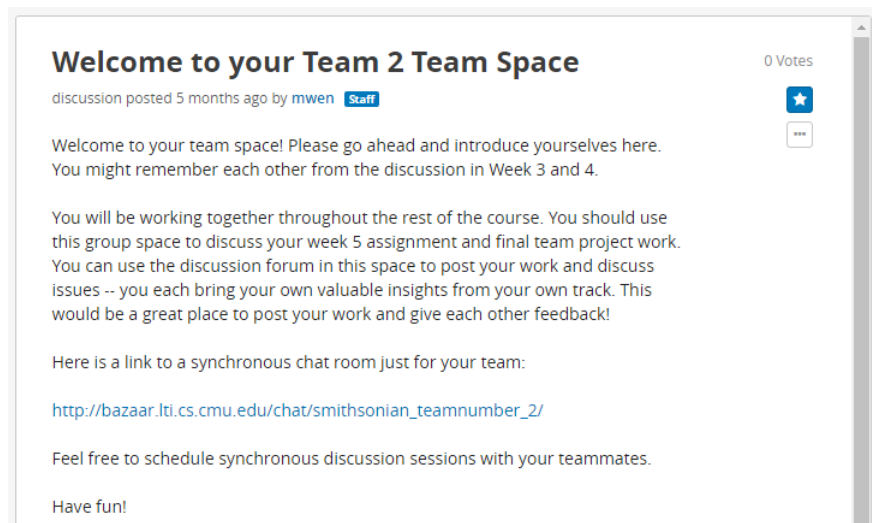


Figure 8.3: Initial post in the team space.

## 8.4.2 Team collaboration communication

In this MOOC, all the teams communicated in the team space. Since there were only 6 teams, we did not observe any team that had problems collaborate. Most teams only tried the synchronous chat in the beginning of the team collaboration. But since it's rare that the team members will log into the chat at the same time. A synchronous tool is not very useful if the team did not set up a meeting time beforehand.

## 8.4.3 Final project submissions

Compared with working alone, what benefit can MOOC student potentially gain from collaborating in teams? In this section, we qualitatively compare the team track submissions to those individual track submissions. We randomly picked five submissions from each track. We summarized and compared all the main components of the project: social issue and archetype; three components of the superhero or superhero team story: build-up, conflict and resolve.

To help students form their superhero story, the project instruction asked students to pick one social issue which the superhero or the superhero team fight for. Individual superheroes seem to focus more on personal issues (e.g. depression and family relationship, Table 8.2). whereas teams seem to focus on bigger societal issues (e.g. equality and nuclear weapons, Table 8.3). Working in a team might raise students' awareness of wider societal problems.

In the project, student need to pick one or more archetypes for their superhero. The most often chosen archetypes in the individual track submissions were Hero and Guardian angel (Table 8.2). In team submissions, the archetypes were usually chosen to complement each other and make the story richer and more interesting, versatile and rich. For example, usually there will be a Leader in a team. Sometimes there will be Healer, Mentor and Anti-hero.

The superhero team story were usually more complicated and included more dialogues. In the build-up of the superhero team stories, they described how these team members formed

**Team Task Description**

In the final two weeks of the course, students collaboratively work on the team project. The team task was split into two parts.

*Course Week 5: Part I Instructions*

For your final project on Track 3: Team Track, you will work in your assigned teams to create a superhero team. To help you do this, we have created the Superhero Team Secret Files

Over the next two weeks, you will work together in your teams to combine your existing superheroes into one team. You'll negotiate the dynamics of the team, create a story involving your team, and write an analysis report explaining your team's actions. We have created the Superhero Team Secret Files to guide you through this work.

*Course Week 6: Part II Instructions*

Last week, you created your team of superheroes by combining your existing superheroes and original superheroes, determining their archetypes, negotiating the team's position on government control vs. personal freedom, and determining the social issue your team will take on. This week, you'll pull together the work you did last week to create your final project.

Your team final project has two parts:

Craft a story that will have your superhero team fighting the social issue you selected, but working within the constraints set by the position that your team took in terms of government control vs. personal freedom. In addition to writing a story, you're also welcome to create two panels as visuals to bring your story to life.

Write a report detailing why your superhero team's selected stance on government control vs. personal freedom is justified considering the characteristic of your superheroes, including the pros and cons of the issue from your superhero team's point of view. Discuss how the team's position about government control vs. personal freedom impacts the team's ability to address the social issue selected.

Figure 8.4: Instructions for the final team project.

the team. Some of them explained where the team got its name. Due to the fact that there were more people (superheroes) involved in the story, the team stories had more conversation between the superheroes. The team project submissions also included more detailed descriptions. For example, there will usually be detailed dialogue about how the superheroes start to know each other and form a team. Or how they decide on tackling a villain together. To endow each superhero with a personality, there are usually scenes where the superheroes have different opinions or don't agree with each other. To bring out these personalities, they incorporated more or humor, especially for those superhero whose archetype is trickster.

<b>Creative Track 1</b>	<b>Summary</b>
Social Issue	Social relations and family issues concerning contact or, lack of it between a parent and a child
Archetypes	Guardian, Angel, Guadian Angel
Build-up	NA
Conflict	a man's ex girlfriend run away with his kid
Resolution	superhero saved the kid though the telegraph
<b>Creative Track 2</b>	<b>Summary</b>
Social Issue	the rising numbers in mental illness
Archetypes	Healer, Protector
Build-up	villain tries to perform something to make the audience mentally ill
Conflict	superman tries to stop the villain who's on the stage performing
Resolution	superhero saved the audience
<b>Creative Track 3</b>	<b>Summary</b>
Social Issue	stress and associated mental health issues
Archetypes	Killer, Witch
Build-up	superman found a victim, the villain sended the victim to fetch some files while setting up an explosion
Conflict	superhero investigates and confront the villain. But if superhero unmake the villain, he'll also unmake himself
Resolution	superhero defeat the villain by passing the victim's pain to the villain
<b>Creative Track 4</b>	<b>Summary</b>
Social Issue	Human trafficking
Archetypes	Hero, Detective
Build-up	superman intestigates international human traffic. a victim has escaped from the ring which the superman was investigating
Conflict	superman goes to the ring, calls police and confront the criminal
Resolution	superhero saved the victims

Table 8.2: Sample creative track projects.

#### 8.4.4 Post-course Survey

The post-course survey examined team collaboration's impact on students' course experience, commitment and satisfaction. There were only 16 students who participated in the team track. In the post course survey, we asked students why they did not participate in the team track. The main reasons were:

1. Lack of time, will not be reliable enough for a team.
2. Prefer to write a solo story of his/her superhero.
3. Worry about the coordination that will be needed for the team track.

For the students who did participate in the team track, they filled in a post-course survey (Appendix B). Most of the students were satisfied with their team experience and team final project submission (4.3 on average, 1-5 range). The difficulties of working in a team include initial communication, time zone issues, team members dropping out or dropping in too late and getting everyone on the same page. Many participants hope to have a better communication tool for the team. Although most teams used the team space in edX as the main communication space, some team chose to communicate using Facebook or email.

On the post-course survey question "I would take this course again to...", 31% of the students indicate that if they take the MOOC again to try the team track. This was the third most frequent reason for taking the MOOC again. Therefore, in the second intervention experiment, we design a team based MOOC as an extension for alumni students who have taken this MOOC before.

### 8.5 Discussion

In this study, we were able to adapt our team formation paradigm in the superhero MOOC. We also gained insights into how to organize the team collaboration in a MOOC. We see that students can collaborate using the beta version of edX team space. Since there were not enough students who participated in the team track, there was no clear evidence for the benefits of our team formation algorithm.

Previous work has already shown that merely forming teams will not significantly improve students' engagement and success [207]. We cannot assume that students will naturally populate the team-based learning in MOOCs: "build it and they will come". Most MOOC students don't yet know why or how they should take advantage of peer learning opportunities. Previous work show that instructors need to take a reinforcing approach online: integrating team-learning systems into the core curriculum and making them a required or extra-credit granting part of the course, rather than optional "hang-out" rooms. In our study, we also observe that students are reluctant to participate in the optional team track since it requires more coordination and work. Therefore, in an educational setting like MOOC, both carrots and sticks are required to keep the commons vibrant [99].

An important step of our team formation paradigm is for the students to first finish individual work before team collaboration. This is to give the team collaboration a momentum so that they don't start with nothing. However, in the context of a MOOC there are less students who actually keep up with the course schedule. Most of the students procrastinate and finish the entire course project in the last week of the course. In this course, there were only 83 students who finished

the individual assignment in Week 1 and 2 by the end of Week 3. Therefore, most students did not have their individual work ready to post in the course forum even if they might want to try the team track. This might be another factor that contributed to the low participation rate in the team track.

On the post-course survey question “I would take this course again to...”, 31% of the students indicate that if they take the MOOC again to try the team track. This was the third most frequent reason for taking the MOOC again. Therefore, in the second intervention experiment, we design a team based MOOC as an extension for alumni students who have taken this MOOC before.



<b>Team Track 1</b>	<b>Summary</b>
Social Issue	Mass murder/violence
Archetypes	Leader + Mentor + Glory-hound + Maverik
Build-up	four supermen found mass murder in separate cities, and identified that it's the same villain
Conflict	a story of how four heroes came to know each other and combine their information
Resolution	superhero team saved the victims
<b>Team Track 2</b>	<b>Summary</b>
Social Issue	defend equality
Archetypes	Leader + Magician + Anti-hero/alien + Action hero+rebel/demi-god
Build-up	Captain Kirk try to recruit his crew members, Great Grandmaster Wizu assigned three members to the team. The mission was to save the Felinans from oppression from DAWG. A process about how each member is on board with the mission, except Hanu.
Conflict	While entering the orbit of the farthest moon of Felinan, the team is surrounded by Felinans, team members fight against the dogs
Resolution	superhero attack villain together, team saved the victims
<b>Team Track 3</b>	<b>Summary</b>
Social Issue	kidnapping
Archetypes	Cheerleader + Helper/Magician + the Intelligent One
build-up	During an interview, Catalyst found the CEO of a child welfare agency might be actually kidnapping. the team started to track the villain
Conflict	the team collaboratively track and locate the villain
Resolution	Catalyst decloak and teleport all the kids out. The team collected more clue about the villain and decided on their next target
<b>Team Track 4</b>	<b>Summary</b>
Social Issue	Fighting against discrimination and oppression of the weak or less fortunate
Archetypes	Leader + Inquisitive + Mentor + Shapeshifter
Build-up	Three girls went missing. Clockwork tracks a villain to a bar by a candy shop
Conflict	Clockwork confronted four burly men and got injured. Siren rushes to distract the villains with his superpower and saved clockwork out. Samsara helped them out. Three superheroes introduced each other.
Resolution	The team realized that the woman that they were trying to save was the actual villain. The story was not finished.
<b>Team Track 5</b>	<b>Summary</b>
Social Issue	ruin the governmental plans to possess a very powerful nuclear weapon
Archetypes	Trickster + Warrior + Magician
Build-up	In one of the small boom towns in California, people live self sufficiently. In recent years there is a multi-year drought, the local farmers have come under increasing pressure to sell their land to a factory farm belonging to a multinational agri-business
Conflict	Poison Ivy has been trying to fight the company but was not successful. Jacque and Lightfoot were tracking some gene-mutated animals. They teamed up and found that the company is doing genetic mutation on animals.
Resolution	Initially thought the company was doing GMO but actually the company was doing genetic mutation on animals.

Table 8.3: Sample team track projects.



## Chapter 9

# Study 7. Team Collaboration as an Extension of the MOOC

In the first intervention experiment, only less than 10% of the entire student population actually participated in the optional team track. However, in the post course survey, 31% of the students indicated that they would like to take the MOOC again to try the team track. And this was the third most frequent reason for taking the same MOOC again. In Study 7, we designed a three week team-based MOOC for alumni of the superhero MOOCs. Therefore, all enrolled students have already finished individual work. To finish this course, all students need to collaborate on a team project of designing a superhero team with the superheroes they designed or analyzed in previous superhero MOOC.

### 9.1 Research Question

Our crowdsourced studies have demonstrated that teams are formed based on transactivity evidence demonstrated better team performance, i.e. knowledge integration. In Study 6, there were not enough students to run the transactivity maximization team formation algorithm, so we did not have enough data from which to gather evidence of how well the team formation algorithm work compare to random team formation. In Study 7, we will try to answer the same question. Although we were not able to do A/B comparison where some teams were formed randomly while other teams were formed based on the algorithm, there was variation in the level of transactive discussion among team members during deliberation. Therefore, the correlation between the level of transactive discussion and the team performance can indicate the success of the team formation method.

Based on the results of our crowdsourced team formation studies (Study 3 and 4), we hypothesis that teams that had more transactive discussions during the community deliberation will have better team process and performance. To measure team performance, we measure the extent to which the superhero team stories are integrated. In particular, we will evaluate (1) how complete the team project is? (2) how many students participated in the project (i.e. how many heroes showed up in the story)? (3) did all the heroes interact with each other in the story? (4) did one of the students (superheroes) dominate the story?

## 9.2 Adapting our team formation paradigm to this MOOC

We slightly modified our team formation paradigm for this mini MOOC. During the actual MOOC, I worked as a course staff. I was responsible for grouping students, sending out team assignment emails, replying to emails and course forum posts that are related to team formation and team collaboration. I served as the “data czar” for SmithsonianX, so I received course data from edX (including enrollment information and discussion posts) weekly.

This MOOC is three weeks long. In the first two course weeks, the instructors prepared new videos and reading material about how to design the essential elements of a superhero team. There was no new instructional material in Week 3.

### 9.2.1 Individual Work

Since all the students in this MOOC have previously finished a course project, so this MOOC did not need course weeks that dedicate to students’ individual work.

### 9.2.2 Course Community Deliberation

In the first week of the course, students participated in course community deliberation. Since all the students enrolled in this course for team collaboration, the course community deliberation was done in the entire course forum. Students need to first post their superhero design or analysis from the original course as a new thread. They will also comment on at least three other peoples heroes. To encourage students to provide feedback transactively, we suggested them to comment on one element of the hero that was successful and on one element where an improvement could be made. At the end of week 1, there were in total 160 students who have posted their previous superhero sketchpad or analysis. There were another 46 students who posted in the second week.

### 9.2.3 Team Formation

In this study, we did the transactive maximization team formation at the end of the first week. We hand annotated 300 reply posts that were randomly sampled from all the discussion posts. Then we trained a similar logistic regression model to predict whether a comment post is transactive or not. Since there were only 10 students who were from the Historical Track, we did not do Jigsaw grouping. Students were assigned to teams of four in the beginning of the second week of the course.

To keep students in a team have similar motivation level, we group verified students and unverified students into teams separately. Verified students had more posts and higher level of transactivity during course community deliberation, as shown in Table 9.1. As a manipulation check to ensure the maximization successfully increased the within team pairwise average transactivity exchange over what would be present from random selection, we compared the number of transactivity interactions with those in randomly formed teams (Table 9.1). The algorithm can significantly increase the average number of transactive discussion in a team.

There were 46 students who posted later than the forum deliberation deadline (then end of week 1). Therefore, 10 teams were formed without evidence of transactivity during week 2.

	Unverified Teams	Verified Teams
Total Teams	37	13
Ave. Transactive Communication	7.81	9.92
Random	0.35	2.08

Table 9.1: Student community deliberation in Week 1.

### 9.2.4 Team Collaboration

The teams collaborated on designing a superhero team in the second and third week of the course. This part is the same as in Study 6. Similar to Study 6, we posted an initial post in each team space. An example is shown in Figure 9.1. In this post, we encouraged students in the team to discuss about their project in this team space. We also encourage them to schedule synchronous meeting. Since the synchronous chat tool was not very useful in Study 6, we did not provide the chat tool for teams in Study 7.

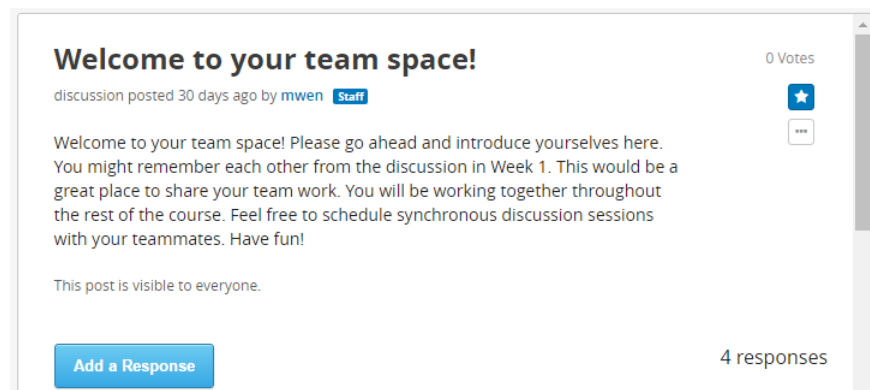


Figure 9.1: Initial Post.

## 9.3 Results

### 9.3.1 Team Collaboration and Communication

#### Stage 1: Self-introductions

Most teams started their collaboration with introducing themselves and their superhero roles. They usually posted their superhero and gave each other further feedback (Figure 9.2).

#### Optional Stage 2: Picking a leader

6 out of the 52 teams explicitly chose a leader in the beginning of their collaboration. An example is shown in Figure 9.3. The team leaders were usually responsible for posting the team shared Google Doc.

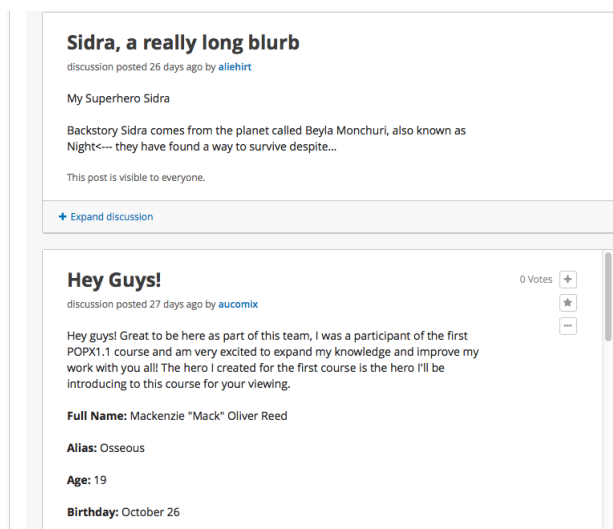


Figure 9.2: Self introductions.

### Stage 3. Collaborating on the team Google Doc

The project submission of this MOOC is in the form of a Google Doc. All the teams shared their superhero team Google Doc in the team space.

Most teams have a team space post for brainstorming the superhero team name and story line. An example is shown in Figure 9.4.

In the post-course survey, almost half of the students indicated that they communicated about once a day with their team.

### 9.3.2 Course Completion

There were in total 770 students who enrolled in this extension team-based MOOC. 106 students have paid for the verified track. Of the 52 formed teams, all the teams submitted their team project at the end of the course.

Out of the 208 students who enrolled in the course, 182 students (87.5%) collaborated in their teams and finished the course. The completion rate in a typical MOOC is 5%. We think there are several factors that contributed to the high retention rate that was observed in Study 7. Our team formation process only groups students who have finished and posted their individual work in the course discussion forum. This screening process ensures that students who are assigned to teams have serious intention of working in a virtual MOOC team. Therefore, it is not surprising that these students have a much higher retention rate. In Study 7, all the enrolled students were alumni from previous offerings of this MOOC, who have already demonstrated effort to finish a similar MOOC. This further increases the retention rate in the extension MOOC. The extension MOOC is short, only three weeks long, compared to a typical 5-6 week MOOC. The carefully designed team-based learning experience, may have also increased students' commitment to the MOOC. It requires further experiments to validate the effect of team-based learning on student commitment in MOOCs.

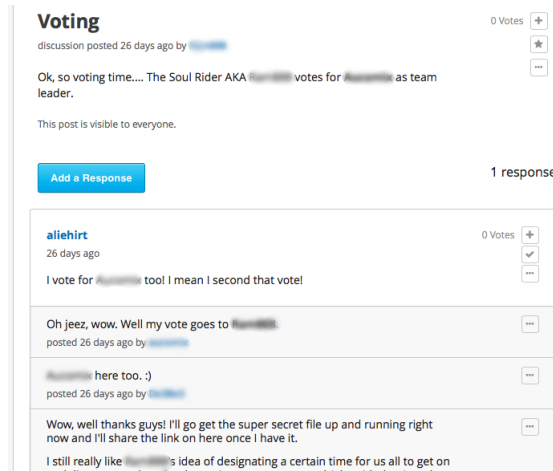


Figure 9.3: Voting for a team leader.

### 9.3.3 Team Project Submission

A SmithsonianX staff who is very familiar with superheroes evaluated all the team projects. She gave each team a 1-4 score for how complete their submission is:

- 4: finished all the components
- 3: only missing panels
- 2: only missing story and panels
- 1: missing more than story or panels

Table 9.2 shows how many teams got each score. Most teams projects, 40 out of 52, contains all the required components.

Score =	4	3	2	1
Total Teams	40	4	7	1

Table 9.2: Finish status of the team projects.

All the teams were formed with the transactivity maximization algorithm. But since teams have various level of amount of transactive discussion during deliberation, we can examine the relationship between the number of transactive discussions among the team members during deliberation and our measures of team process/performance. Controlling for whether the team members were verified students, and whether there was one historical track student in the team, the number of transactive communication among team members during the community deliberation had a marginal effect on the finish status score ( $F(3, 48)=1.83, p = 0.1$ ). The verified status of the team ( $p = 0.18$ ) and whether it was a mixed team ( $p = 0.47$ ) had no significant effect on the finish status of the team project. Since our teams were formed where teams' overall transactivity communication were maximized, this indicate that our team formation method can improve overall team performance, even in the real MOOC context where many factors influence team collaboration.

To examine the effect of grouping on team collaboration participation, we counted how

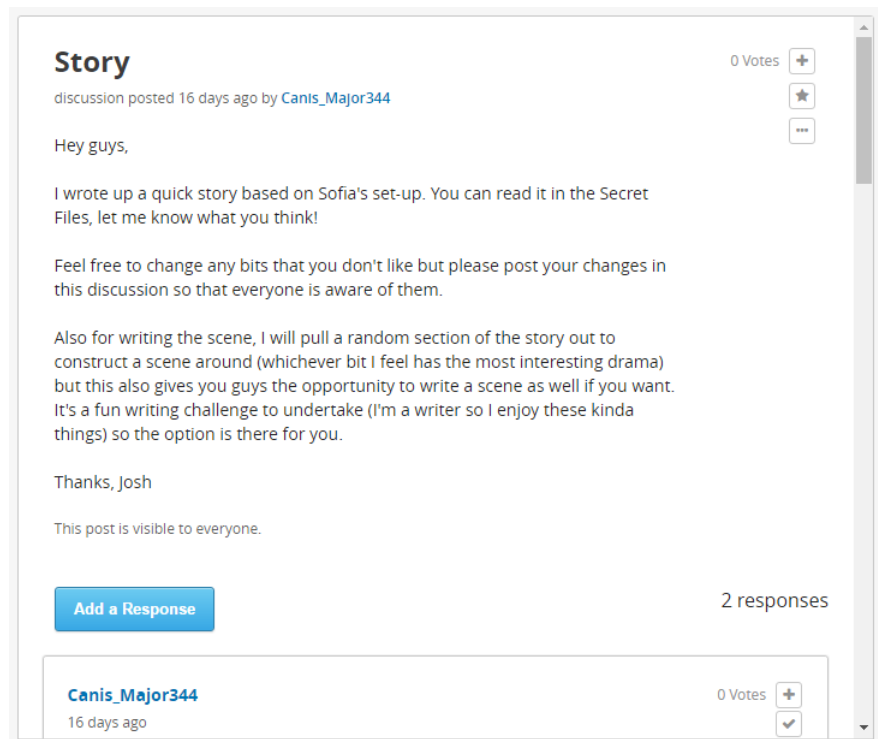


Figure 9.4: A brainstorming post.

many students actually participate in the team project. For each student who participated, their hero/villain needs to show up in the story. Controlling for whether the team members were verified students, and whether there was one historical track student in the team, the number of transactive communication among team members during the community deliberation had a significant effect on how many students participated in the collaboration ( $F(3,48) = 1.90$ ,  $p = 0.02$ ). The verified status of the team ( $p = 0.89$ ) and whether it was a mixed team ( $p = 0.81$ ) had no significant effect on how many students participated in the collaboration.

As a reflection of whether students in the team interacted with each other, we check if all of the superheroes have interacted with each other in the story. Overall, in 44 out of 52 superhero team stories, there was at least one scene where all four superheroes interacted with each other. The number of transactive communication during the community deliberation had a marginal effect on whether all the superheroes have interacted or not ( $F(3,48) = 3.02$ ,  $p = 0.066$ ). The verified status of the team had a marginal effect on whether all the superheroes have interacted or not ( $F(3,48) = 3.02$ ,  $p = 0.04$ ). Superheroes in a verified team was significantly more likely to have interacted with each other.

There were 15 teams where there was one student (superhero) who dominated the superhero team story. The number of transactive communication during the community deliberation has no significant effect on whether there was one superhero that dominates the story ( $F(3,41) = 1.50$ ,  $p = 0.80$ ). The verified status of the team had no significant effect on whether all the superheroes have interacted or not ( $p = 0.59$ ). A mixed team where there was one student from the Historical Track was more likely to have one superhero that dominates the story ( $p = 0.05$ ).



How often did you communicate with your team? (137 responses)

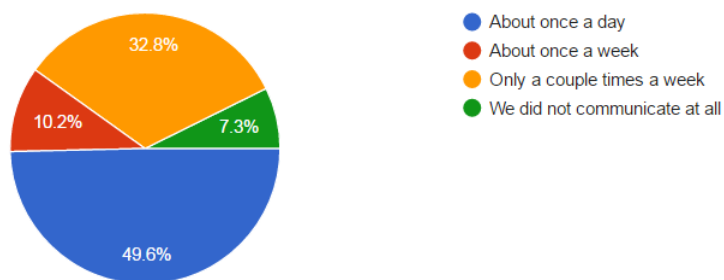


Figure 9.5: Team communication frequency survey.

### Observations on how transactivity maximization team formation affected the teams

Since many teams chose to communicate outside the team space, the total number of posts in the team space is not representative of how much discussion the team had during collaboration. We qualitatively read all the posts in the team spaces. Since the transactivity maximization team formation tends to match students with whom they have transactively discussed during deliberation, many students recognized their team members with whom they have interacted with during community deliberation: "I've already read your story in week 1", "I am happy to see that my team members had characters that I was familiar with." "Sup Ron, first of all; thanks for your comments on Soldier Zeta". This created a friendly start for the team. Some students indicated that they already had idea about the superhero team story: "I can already see Osseus and the Soul Rider bonding over the fact your character had a serious illness and The Soul Rider brother was mentally handicapped", "We've already exchanged some ideas last week, I think we can have some really fun dynamics with our crew of heroes!".

### 9.3.4 Survey Results

There were in total 136 students who responded to our optional, anonymous post-course survey. Satisfaction with the team experience was rated on a scale from 1-5 with 5 being very satisfied. On average, the satisfaction score is 4.20 out of 5. Satisfaction with the project your team turned in was also rated on a scale from 1-5 with 5 being very satisfied. On average, the satisfaction score is 3.96 out of 5. Overall, students are satisfied with their team experience and project submission.

The end of program feedback asked the question, "What was most difficult about working in your team?" The responses here had two main themes: Time zones and the Team Discussion Space.

(1) Time zones: Much of the feedback from students was that it was very difficult to communicate with their teammates because of times zones. They had trouble finding times to chat live and also found it difficult to agree on changes or make progress since either changes were made while some team members were offline or it took so long to make a decision the project felt rushed.

(2) Team Discussion Space: Since there was no support for synchronous communication and personal messaging functions in edX, many teams rely on the team space as the main team communication tool. This is not very efficient way of communication. Most of the teams communicated asynchronously over the chat and comment functions in the Google Doc. 10 teams scheduled synchronous meeting over Skype/Google Hangout in the team space. At least 6 teams communicated using Facebook groups. In the post course survey, 20% of the students indicated that they used email to communicate with their team members.

Many students felt that the team discussion space was difficult to use in the sense that messages got buried and students were not notified when there was a new message in their team space. This made keeping on top of the discussion challenging unless students remembered to check each day and put in the effort to sort through the messages. We think a well-integrated team synchronous communication and a messaging tool will be helpful to include in the team space.

## **9.4 Discussions**

### **9.4.1 Problems with team space**

The team space module we used is a beta version. The list of teams and team space discussions were visible to all students who enrolled in the MOOC (Figure 9.6). All the teams relied primarily on the team space as their team management tool. A screenshot of a team space is shown in Figure 9.7. It does not allow instructors to limit who can join the team. So anyone can join a team that's not full yet. After team formation, there were three students who did not post to the forum on time and took the spot of the others. The original team members had to email us for help. We had to contact those students and ask them to join a new team. This problem can be solved by adding a verification to each team where only assigned team members can join the team.

### **Coordinating Team Collaboration**

In this deployment study, our main intervention was team formation. We did not provide structured team collaboration support to the teams. One course staff checked each team space daily and provided feedback to the team work. In the course, we asked students who had some trouble getting their team going or working in their team to email us. We offered Skype/Google Hangout team mentoring to these students. There were two teams where conflicts happened. We offered to hold a Skype meeting to reconcile the miscommunication. However, they refuse to work with each other anymore. In the end, we had to split two of the problematic teams into two teams. When tension has already built up, people may become more reluctant to meet with each other and resolve the issue. The problems these teams encountered are mostly due to that team members were on different pace with the course. When some of them joined the teamwork late, they felt difficult to contribute as much as they would like to. Or the other team members were not willing to incorporate the late-comers' edits or suggestions into their work. We think clearer guidelines about working in a MOOC teams may help these teams.

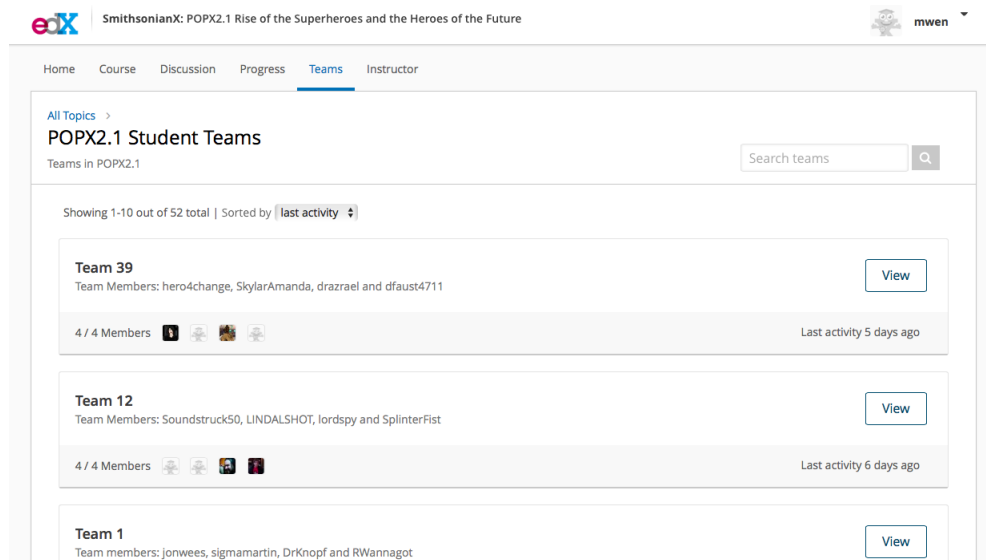


Figure 9.6: Team space list.

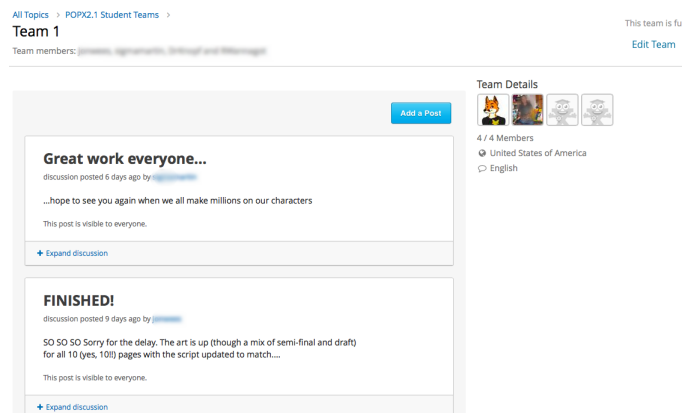


Figure 9.7: Team space in edX.

How to address the problem where students in a MOOC team usually have their own schedule is an open question. We think it is important to have a “kick-off” meeting at the beginning of the team collaboration, which may help synchronous team members and establish a team collaboration schedule. Another possible way, is to split the team task into modules and send out reminders before the expected deadline of each module. A few teams adopted a to-do list to split the project in to smaller tasks. An example is shown in Figure 9.8. To-do list seems to be an effective way of coordinating virtual team members. Kittur et al. has explored methods for task decomposition as a means of achieving complex work products [95].

## Benefit of Community Deliberation

One benefit of course community deliberation is students can potentially get feedback and support from all the students in the course. In week 1 of the MOOC, 208 students posted more than

**To Do List 7/16-7/23**

question posted 24 days ago by [cjh0215](#)

2 Votes

I took the liberty to make our to-do list. If you can finish well ahead of the date go right ahead! If you want to swap a role with someone else, let them know. While we have so many great ideas, I think it would be best if we can come up with a short story, kinda like a one-shot or a pilot episode. We're in a bit of a time crunch, the Super Hero Secret Files only have a few slides (we could duplicate a few), and I'm sure we're busy with work and personal life. This is fun! I really do love the amount of creativity in this team!

**Due Tuesday night 7/19**

1. Everyone fill out your superhero's bio on the superhero secret files. I added each hero's name in alphabetical order. I will draw each hero.
2. Everyone brainstorm and decide our social or political issue with sources/articles to back it up (slide 10). Decide by Tuesday night.
3. Everyone brainstorm 1 new super villain or super villain team. (slide 11...but if we go by team we could duplicate the slides)- *Alias, Super Power, and Archetype* Decide by Tuesday night.
4. Everyone brainstorm Construct Your Story (slide 12) *Build Up, Conflict, Resolution* Decide by Tuesday night.

**Due Wednesday night 7/20**

- Crystal - Add the hero bio images
- MarkH1990 - Add "Build Up" paragraph in Construct Your Story (slide 12)

Figure 9.8: A to-do list.

one thousand posts and comments about their superhero designs. Although we did not ask the students to discuss transactively, 60% of them were transactive discussion. At the end of the course, 44 of the 52 teams voluntarily posted their superhero team story to the forum for further community feedback. The discussion and feedback in the course forum continued even after the course has ended. We think this demonstrate the benefit of having course community discussion before small team collaboration.

### Team collaboration as an extension MOOC

In Study 6, we explored one possible setting of team-based learning in MOOC, i.e., provide an optional team track of study. There was very low participation in the team track. In Study 7, many students successfully finished an extension team-based MOOC where all students work in small teams to finish the course. From our two deployment studies, we see that more students participated team-based MOOC setting where team collaboration is mandatory and better incorporated into the MOOC. We also received emails from students inquiring about when will the team extension MOOC be offered again. Due to the popularity of this extension MOOC, SmithsonianX plan to offer it again in the future.

The reasons why a team-based MOOC is more popular than an optional team track may include: (1) Compared to individual track, team track study requires more work and coordination. Students need more incentives to participate in team collaboration in MOOC if it is not required to pass the course. (2) In a team-based MOOC where everyone works in team collaboration, the instruction materials may adapt to the team project or team collaboration. In Study

7, the video lectures contained information about how to design a superhero team, which is the team project.(3) Our team formation paradigm requires students to finish individual work before small team collaboration. Less MOOC students were able to finish their individual work by the deadline.

## **Different tracks**

Across the two deployment studies, most of the participants in the team collaboration are students from the Creative Track. Out of the 208 people who participated in this MOOC, only 10 of them are from the Historical Track. We formed 10 mixed groups that have three Creative track students and one Historical Track student. We did not observe significant impact of having a team member from the Historical Track in terms of complete status, interaction and number of students who participated. However, teams with a Historical Track team member are significantly more likely to have one student (superhero) who dominates the story.

### **9.4.2 Team Member Dropout**

Team member dropout is one of the challenges coordinating team-based MOOCs. Although in this deployment study, all the teams had at least two members who were active till the end of the course. We did receive two emails from the students earlier in the course which told us there were no other team member that was active in their team. Their team members did end up join them later in the course. If there was only one active student left in a team, we need to either re-group students if all their team members drop out. Or the course need to design an individual version of the course project so that the student could finish the course without collaborators.

### **9.4.3 The Role of Social Learning in MOOCs**

Online learning should have a social component ; how to do that in the context of Massive Open Online Courses (MOOCs) is an open question. This research addresses the question of how to improve the online learning experience in contexts in which students can collaborate with a small team in a MOOC. Forums are pervasive in MOOCs and have been characterized as an essential ingredient of an effective online course [118], but early investigations of MOOC forums show struggles to retain users over time. As an alternative form of engagement, informal groups associated with MOOC courses tend to spring up around the world. Students use social media to organize groups or decide to meet in a physical location to take the course together.

In classrooms, team-based learning has been widely used in social social science classes as well as natural science class such as math or chemistry [173]. Most of the MOOCs that adopted the "talkabout" MOOC group discussion tool were social science MOOCs<sup>1</sup>. We think that social science MOOCs would benefit the most from having a team-based learning component.

<sup>1</sup><https://talkabout.stanford.edu/welcome>

## 9.5 Conclusion

In Study 7, we were able to adapt our team formation paradigm in a three-week extension MOOC. Similar to Study 6, we see that most teams can collaborate and finish their team project using the beta version of edX team space. We also gained insights into how to better coordinate and support team collaboration in a MOOC.

Despite the contextual factors that might influence the results, we see that the teams that had higher level of transactivity during deliberation had more students who participated in the team project and better team performance. This correlational result combined with the causal results in the controlled MTurk studies demonstrate the effectiveness of our team formation process. In future, it will be interesting to do A/B testing on the team formation method and compare algorithm assigned teams with self-selected teams.

# Chapter 10

## Conclusions

The main goal of my dissertation is to support team-based learning, especially in MOOC environments where students have different background and perspectives. To support virtual team formation, I started by investigating the type of process measures that are predicative of student commitment and team performance (Study 1 and 2).

Based on the corpus analysis in Study 1 and 2, we designed a deliberation-based team formation where students hold a course community deliberation before small group collaboration. The key idea is that students should have the opportunity to interact meaningfully with the community before assignment into teams. That discussion not only provides evidence of which students would work well together, but it also provides students with a wealth of insight into alternative task-relevant perspectives to take with them into the collaboration. Then we form teams that already demonstrate good team process, i.e. transactive discussions, during course community deliberation. Next, in Study 3 and 4, we evaluated this virtual team formation process in crowdsourced environments using the team process and performance as success measures. Ultimately, my vision was to automatically support team-based learning in MOOCs. In this dissertation, I have presented methods for automatically support team formation and team collaboration discussion (Study 3 - 5). Given my initial success in the crowdsourced environment, the final two studies examine the external validity within two real MOOCs with different team-based learning settings.

### 10.1 Reflections on the use of MTurk as a proxy for MOOCs

Rather than trying out untested designs on real live courses, we have prototyped and tested the approach using a crowdsourcing service, Amazon Mechanical Turk. MTurk is extremely useful for piloting designs. It is much faster to do MTurk experiments than do lab studies and deployment studies. We were able to quickly iterate on our designs. In our crowdsourced studies, we have experimented with over 200 teams, almost 800 individual Turkers. It is very difficult to recruit such a big crowd for lab studies.

While crowd workers likely have different motivations from MOOC students, their remote individual work setting without peer contact resembles today's MOOC setting where most students learn in isolation [41]. MTurk experiments gave us confidence that the crowdsourced work-

flow will also work for MOOC students. With adaptation, we were able to use the same team formation process in our MTurk experiments in real MOOCs. We think MTurk experiments are appropriate for testing out team collaboration/communication support designs for other real-world setting where participants are virtually communicating with each other, such as online community setting.

Similar to lab studies, crowdsourced environments are appropriate for controlled experiments that test the internal validity of the hypothesis. Under the course or instructor's constraints, it is difficult to do A/B testing in real MOOCs.

Crowdsourced experiments may not represent how MOOC students will adopt or enjoy the designs. Since crowd workers did the task to earn money, the main factor that affect crowd workers' intention to take or finish the task is how much money is paid. This is obviously not the case for MOOC students. It is crucial to understand how many students will actually adopt or enjoy the designs by doing deployment studies.

## **10.2 Contributions**

My dissertation work makes contributions to the fields of online learning, Computer-Supported Collaborative Learning and conversational analysis. The contributions span theoretical, methodological, and practical arenas. Building on a learning science concept of Transactivity, we designed and validated a team formation paradigm for team-based learning in MOOCs. My thesis includes practical design advice and team coordination tools for MOOC practitioners.

## **10.3 Future Directions**

This thesis opens up several avenues of possible exploration for many different features of team-based learning or team collaboration that are being adapted to online contexts. This includes different types of individual work that could prepare participants for their team collaboration, ways of supporting online community deliberation, and support for team communication and collaboration in MOOCs.

### **10.3.1 Compare with self selection based team formation**

In this thesis, we have compared transactivity based team formation with random team formation in crowdsourced environment. In future work, we would like to compare our team formation method with self-selection based team formation, both in the crowdsourced environment and in real MOOCs. With the

### **10.3.2 Team Collaboration Support**

In future work, we would like to explore how to better support team collaboration, especially in future MOOC deployment studies. In our deployment study, several teams had conflicts during



collaboration. They may benefit from team collaboration support that helps the team members stay on the same page.

### **10.3.3 Community Deliberation Support**

The number of transactive discussions between team members were correlated with team performance. In future work, we can provide more scaffolding to increase the level of transactivity during community deliberation. In both our MTurk and deployment studies, in order to encourage participants to discuss transactively, we designed discussion instructions as scaffolding. In future work, we would like to explore providing dynamic deliberation support. For example, we can provide real-time feedback to the student on how to make his/her post more transactive. During big community deliberation, it is challenging for students to browse through hundreds or even thousands of threads. When there are too many participants in the community deliberation (e.g. more than 1000), navigability may become another challenge for participants' online community deliberation [177]. We can recommend threads to a participant when he joins the community deliberation based on his individual work. These deliberation tool may provide scaffolding for transactive discussion during community deliberation.

### **10.3.4 Team-based Learning for Science MOOC**

The team formation process in this thesis is tested within a social science MOOC. Whether it will work for a programming or math MOOC is an open question. Imaginably, it will be challenging to design a community deliberation task that is valuable for a programming or math MOOC.



## Chapter 11

### APPENDIX A: Final group outcome evaluation for Study 4

We evaluate the final group outcome based on the scoring rubrics shown in Table 11.1- 11.4.

Score	Energy source X Requirement	Example
1	The sentence explicitly, correctly considers at least one of the eight requirements for one energy source.	(1) While windfarms may be negative for the bird population, we would have to have environmental studies done to determine the impact. (2) Hydro power is an acceptable solution to sustainable energy but the costs may outweigh the benefits.
0	The sentence does not explicitly consider one of the eight requirement for one energy source, or the statement is wrong.	(1) Hydro does not release any chemicals into the water. Reason: this is not one of the eight requirements (2) Nuclear power other than safety it meets all requirements. Reason: this statement is wrong since nuclear power does not meet all the other requirements. (3) I agree, wind power is the way to go. Reason: this statement did not explicitly consider one of the eight requirements.

Table 11.1: Energy source X Requirement

Score	Energy X Energy Comparison	Example
1	The sentence explicitly, compares two energy resources.	It is much more expensive to build a hydro power plant than it is to run a windfarm and and the city has a tight energy budget. Reason: this statement compares hydro power and wind power.
0	The sentence does not explicitly compares two energy resources.	The drawback is that wind may not be reliable and the city would need to save up for a better source Reason: this statement only refers to one energy source.

Table 11.2: Energy X Energy Comparison

Score	Address additional valid requirements	Example
1	The statement correctly considers one or more requirement that is not one of the eight requirements.	(1)Hydro plants also require a large body of water with a strong current which may and the city has a tight energy budget. not be available in the city. Reason: this statement considers a requirement of a large body of water for hydro power to work for this city. But this requirement is not one of the eight requirements. (2) Coal - Abundant, cheap, versatile Reason: Abundant and versatile are extra requirements.
0	The sentence does not consider an extra requirement or the statement is wrong.	(1)The wind farm could be placed on the large amounts of farmland, as there is adequate space there. Reason: farmland is one of the requirements energy source.

Table 11.3: Address additional valid requirements.

Score	Incorrect requirement statements	Example
1	The sentence contains wrong statement or the sentence is overly conclusive.	(1)Hydro does not affect animal life around it. Reason: hydro power has a big impact on the natural environment nearby. not be available in the city. Reason: this statement considers a requirement of a large body of water for hydro power to work for this city. But this requirement is not one of the eight requirements. (2) Coal - Abundant, cheap, versatile Reason: Abundant and versatile are extra requirements.
0	The statement in the sentence is mostly correct.	(1) Hydro does not affect animal life around it. Reason: hydro power has a big impact on the natural environment nearby.

Table 11.4: Address additional valid requirements.



## Chapter 12

# APPENDIX B. Survey for Team Track Students

1. What did you expect to gain from joining a team?
  - A. Get feedback or help
  - B. Make the course more social
  - C. Take on a more challenging project
  - D. Other, please specify:
  
2. How satisfied were you with your team's experience overall?  
Not at all satisfied -¿ very satisfied
  
3. How satisfied were you with the project your team turned in if any?  
Not at all satisfied -¿ very satisfied
  
4. What was the biggest benefit of participation in the team?
  - A. Get feedback or help
  - B. Make the course more social
  - C. Take on a more challenging project
  - D. Other, please specify:
  
5. What was missing most from your team experience?
  - A. Get feedback or help
  - B. Make the course more social
  - C. Take on a more challenging project
  - D. Other, please specify:
  
6. How often did you communicate with your team?
  - A. About once a day
  - B. About once a week
  - C. Only a couple of times
  - D. We did not communicate at all

7. What technologies did you use for your team work? Please check all that apply.

- a. EdX team space
- b. EdX messages
- c. EdX discussion board
- d. Chat tool
- e. Email
- f. Smithsonian discussion board
- g. Skype or Google Hangout
- h. Phone call
- i. Other tool, please specify:

8. What was most difficult about working in your team?

9. Which aspects would you have preferred to be managed differently (for example, how and when teams were assigned, the software infrastructure for coordinating team work, tools for communication?) Please make specific suggestions.

10. If you take another MOOC in the future with a team track option, would you choose to take the team track? Why or why not?



# Bibliography

- [1] David Adamson. *Agile Facilitation for Collaborative Learning*. PhD thesis, Carnegie Mellon University Pittsburgh, PA, 2014. 2.3.2, 7, 7.1
- [2] David Adamson, Colin Ashe, Hyeju Jang, David Yaron, and Carolyn P Rosé. Intensification of group knowledge exchange with academically productive talk agents. In *The Computer Supported Collaborative Learning (CSCW) Conference 2013, Volume 1*, page 10. Lulu. com. 2.4
- [3] David Adamson, Gregory Dyke, Hyeju Jang, and Carolyn Penstein Rosé. Towards an agile approach to adapting dynamic collaboration support to student needs. *International Journal of Artificial Intelligence in Education*, 24(1):92–124, 2014. 4.4.3
- [4] Rakesh Agrawal, Behzad Golshan, and Evimaria Terzi. Grouping students in educational settings. In *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 1017–1026. ACM, 2014. 2.3.1
- [5] Ravindra K Ahuja and James B Orlin. A fast and simple algorithm for the maximum flow problem. *Operations Research*, 37(5):748–759, 1989. 3.3.4, 6.4
- [6] Hua Ai, Rohit Kumar, Dong Nguyen, Amrut Nagasunder, and Carolyn P Rosé. Exploring the effectiveness of social capabilities and goal alignment in computer supported collaborative learning. In *Intelligent Tutoring Systems*, pages 134–143. Springer, 2010. 2.2, 6.2.1
- [7] Carlos Alario-Hoyos. Analysing the impact of built-in and external social tools in a mooc on educational technologies. In *Scaling up learning for sustained impact*, pages 5–18. Springer, 2013. 2.1.2
- [8] Enrique Alfonseca, Rosa M Carro, Estefanía Martín, Alvaro Ortigosa, and Pedro Paredes. The impact of learning styles on student grouping for collaborative learning: a case study. *User Modeling and User-Adapted Interaction*, 16(3-4):377–401, 2006. 2.3.1
- [9] Aris Anagnostopoulos, Luca Becchetti, Carlos Castillo, Aristides Gionis, and Stefano Leonardi. Online team formation in social networks. In *Proceedings of the 21st international conference on World Wide Web*, pages 839–848. ACM, 2012. 2.3.1
- [10] Antonio R Anaya and Jesus G Boticario. Clustering learners according to their collaboration. In *Computer Supported Cooperative Work in Design, 2009. CSCWD 2009. 13th International Conference on*, pages 540–545. IEEE, 2009. 2.3.1, 2.3.2
- [11] Elliot Aronson. *The jigsaw classroom*. Sage, 1978. 6.1, 6.2.1

- [12] Margarita Azmitia and Ryan Montgomery. Friendship, transactive dialogues, and the development of scientific reasoning. *Social development*, 2(3):202–221, 1993. 2.2
- [13] Lars Backstrom, Dan Huttenlocher, Jon Kleinberg, and Xiangyang Lan. Group formation in large social networks: membership, growth, and evolution. In *Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 44–54. ACM, 2006. 2.3.1
- [14] Donald R Bacon, Kim A Stewart, and William S Silver. Lessons from the best and worst student team experiences: How a teacher can make the difference. *Journal of Management Education*, 23(5):467–488, 1999. 2.3.1
- [15] Michael Baker and Kristine Lund. Promoting reflective interactions in a cscl environment. *Journal of computer assisted learning*, 13(3):175–193, 1997. 2.1.4
- [16] Sasha A Barab and Thomas Duffy. From practice fields to communities of practice. *Theoretical foundations of learning environments*, 1(1):25–55, 2000. 3.2.1
- [17] Brigid Barron. When smart groups fail. *The journal of the learning sciences*, 12(3):307–359, 2003. 2.2
- [18] James Bo Begole, John C Tang, Randall B Smith, and Nicole Yankelovich. Work rhythms: analyzing visualizations of awareness histories of distributed groups. In *Proceedings of the 2002 ACM conference on Computer supported cooperative work*, pages 334–343. ACM, 2002. 2.3.2
- [19] Sue Bennett, Ann-Marie Priest, and Colin Macpherson. Learning about online learning: An approach to staff development for university teachers. *Australasian Journal of Educational Technology*, 15(3):207–221, 1999. 2.1.2
- [20] Marvin W Berkowitz and John C Gibbs. Measuring the developmental features of moral discussion. *Merrill-Palmer Quarterly (1982-)*, pages 399–410, 1983. (document), 0.1, 1, 3.2.1, 6.1
- [21] Camiel J Beukeboom, Martin Tanis, and Ivar E Vermeulen. The language of extraversion extraverted people talk more abstractly, introverts are more concrete. *Journal of Language and Social Psychology*, 32(2):191–201, 2013. 4.3.3, 4.3.3
- [22] Kenneth A Bollen. Total, direct, and indirect effects in structural equation models. *Sociological methodology*, pages 37–69, 1987. 3.2.2
- [23] Richard K Boohar and William J Seiler. Speech communication anxiety: An impediment to academic achievement in the university classroom. *The Journal of Classroom Interaction*, pages 23–27, 1982. 2.1.2
- [24] Marcela Borge, Craig H Ganoë, Shin-I Shih, and John M Carroll. Patterns of team processes and breakdowns in information analysis tasks. In *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work*, pages 1105–1114. ACM, 2012. 5.4.1
- [25] Michael T Brannick and Carolyn Prince. An overview of team performance measurement. *Team performance assessment and measurement. Theory, methods, and applications*, pages 3–16, 1997. 5.4

- [26] Christopher G Brinton, Mung Chiang, Shaili Jain, Henry Lam, Zhenming Liu, and Felix Ming Fai Wong. Learning about social learning in moocs: From statistical analysis to generative model. *arXiv preprint arXiv:1312.2159*, 2013. 2.1.3
- [27] Christopher G Brinton, Mung Chiang, Sonal Jain, HK Lam, Zhenming Liu, and Felix Ming Fai Wong. Learning about social learning in moocs: From statistical analysis to generative model. *Learning Technologies, IEEE Transactions on*, 7(4):346–359, 2014. 4
- [28] Robert M Carini, George D Kuh, and Stephen P Klein. Student engagement and student learning: Testing the linkages\*. *Research in higher education*, 47(1):1–32, 2006. 4.1
- [29] Robert J Cavalier. *Approaching deliberative democracy: Theory and practice*. Carnegie Mellon University Press, 2011. 1, 6.1
- [30] Justin Cheng, Chinmay Kulkarni, and Scott Klemmer. Tools for predicting drop-off in large online classes. In *Proceedings of the 2013 conference on Computer supported co-operative work companion*, pages 121–124. ACM, 2013. 4.3.1
- [31] Michelene TH Chi. Quantifying qualitative analyses of verbal data: A practical guide. *The journal of the learning sciences*, 6(3):271–315, 1997. 3.2.1
- [32] Michelene TH Chi. Self-explaining expository texts: The dual processes of generating inferences and repairing mental models. *Advances in instructional psychology*, 5:161–238, 2000. 3.2.1
- [33] Michelene TH Chi. Active-constructive-interactive: A conceptual framework for differentiating learning activities. *Topics in Cognitive Science*, 1(1):73–105, 2009. 3.2.1
- [34] Michelene TH Chi and Miriam Bassok. Learning from examples via self-explanations. *Knowing, learning, and instruction: Essays in honor of Robert Glaser*, pages 251–282, 1989. 3.2.1
- [35] Michelene TH Chi, Stephanie A Siler, Heisawn Jeong, Takashi Yamauchi, and Robert G Hausmann. Learning from human tutoring. *Cognitive Science*, 25(4):471–533, 2001. 3.2.1
- [36] Lydia B Chilton, Clayton T Sims, Max Goldman, Greg Little, and Robert C Miller. Seaweed: A web application for designing economic games. In *Proceedings of the ACM SIGKDD workshop on human computation*, pages 34–35. ACM, 2009. 3.3.2
- [37] Gayle Christensen, Andrew Steinmetz, Brandon Alcorn, Amy Bennett, Deirdre Woods, and Ezekiel J Emanuel. The mooc phenomenon: who takes massive open online courses and why? *Available at SSRN 2350964*, 2013. 4.3.1
- [38] Christos E Christodoulopoulos and K Papanikolaou. Investigation of group formation using low complexity algorithms. In *Proc. of PING Workshop*, pages 57–60, 2007. 2.3.1
- [39] S Clarke, Gaowei Chen, K Stainton, Sandra Katz, J Greeno, L Resnick, H Howley, David Adamson, and CP Rosé. The impact of cscl beyond the online environment. In *Proceedings of Computer Supported Collaborative Learning*, 2013. 7.1
- [40] Doug Clow. Moocs and the funnel of participation. In *Proceedings of the Third International Conference on Learning Analytics and Knowledge*, pages 185–189. ACM, 2013. 2.1.2

- [41] D Coetzee, Seongtaek Lim, Armando Fox, Björn Hartmann, and Marti A Hearst. Structuring interactions for large-scale synchronous peer learning. In *CSCW*, 2015. 1, 3.3.3, 6.1, 6.2.1, 6.3, 10.1
- [42] Derrick Coetzee, Armando Fox, Marti A Hearst, and Bjoern Hartmann. Chatrooms in moocs: all talk and no action. In *Proceedings of the first ACM conference on Learning@scale conference*, pages 127–136. ACM, 2014. 2.1.3
- [43] Elisabeth G Cohen and Rachel A Lotan. *Designing Groupwork: Strategies for the Heterogeneous Classroom Third Edition*. Teachers College Press, 2014. 2.3.1
- [44] Elizabeth G Cohen. Restructuring the classroom: Conditions for productive small groups. *Review of educational research*, 64(1):1–35, 1994. 3.2.1
- [45] James S Coleman. *Foundations of social theory*. Harvard university press, 1994. 2.3.1
- [46] Lyn Corno and Ellen B Mandinach. The role of cognitive engagement in classroom learning and motivation. *Educational psychologist*, 18(2):88–108, 1983. 3.2.1, 4.4.3
- [47] Stata Corp. *Stata Statistical Software: Statistics; Data Management; Graphics*. Stata Press, 1997. 3.2.2
- [48] Catherine H Crouch and Eric Mazur. Peer instruction: Ten years of experience and results. *American journal of physics*, 69(9):970–977, 2001. 2.1.4
- [49] Cristian Danescu-Niculescu-Mizil, Moritz Sudhof, Dan Jurafsky, Jure Leskovec, and Christopher Potts. A computational approach to politeness with application to social factors. *ACL*, 2013. 5.4.2
- [50] Eustáquio São José de Faria, Juan Manuel Adán-Coello, and Keiji Yamanaka. Forming groups for collaborative learning in introductory computer programming courses based on students’ programming styles: an empirical study. In *Frontiers in Education Conference, 36th Annual*, pages 6–11. IEEE, 2006. 2.3.1
- [51] Richard De Lisi and Susan L Golbeck. Implications of piagetian theory for peer learning. 1999. 1, 2.1.3, 2.2, 2.3.2, 3.2.1, 6.1, 6.4.2
- [52] Bram De Wever, Tammy Schellens, Martin Valcke, and Hilde Van Keer. Content analysis schemes to analyze transcripts of online asynchronous discussion groups: A review. *Computers & Education*, 46(1):6–28, 2006. 3.2.1
- [53] Jennifer DeBoer, Glenda S Stump, Daniel Seaton, and Lori Breslow. Diversity in mooc students backgrounds and behaviors in relationship to performance in 6.002 x. In *Proceedings of the Sixth Learning International Networks Consortium Conference*, 2013. 4.1
- [54] Ronald Decker. Management team formation for large scale simulations. *Developments in Business Simulation and Experiential Learning*, 22, 1995. 2.3.1
- [55] Katherine Deibel. Team formation methods for increasing interaction during in-class group work. In *ACM SIGCSE Bulletin*, volume 37, pages 291–295. ACM, 2005. 2.3.1
- [56] Scott DeRue, Jennifer Nahrgang, NED Wellman, and Stephen Humphrey. Trait and behavioral theories of leadership: An integration and meta-analytic test of their relative validity. *Personnel Psychology*, 64(1):7–52, 2011. 5.4.3

- [57] Willem Doise, Gabriel Mugny, A St James, Nicholas Emler, and D Mackie. *The social development of the intellect*, volume 10. Elsevier, 2013. 2.3.1
- [58] Martin Dowson and Dennis M McInerney. What do students say about their motivational goals?: Towards a more complex and dynamic perspective on student motivation. *Contemporary educational psychology*, 28(1):91–113, 2003. 4.3.2
- [59] Gregory Dyke, David Adamson, Iris Howley, and Carolyn Penstein Rosé. Enhancing scientific reasoning and discussion with conversational agents. *Learning Technologies, IEEE Transactions on*, 6(3):240–247, 2013. 2.3.2
- [60] Kate Ehrlich and Marcelo Cataldo. The communication patterns of technical leaders: impact on product development team performance. In *CSCW*, pages 733–744. ACM, 2014. 5.4.3
- [61] Richard M Emerson. Social exchange theory. *Annual review of sociology*, pages 335–362, 1976. 2.3.1
- [62] Timm J Esque and Joel McCausland. Taking ownership for transfer: A management development case study. *Performance Improvement Quarterly*, 10(2):116–133, 1997. 4.3.2
- [63] Patrick J Fahy, Gail Crawford, and Mohamed Ally. Patterns of interaction in a computer conference transcript. *The International Review of Research in Open and Distributed Learning*, 2(1), 2001. 4.3.3
- [64] Rong-En Fan, Kai-Wei Chang, Cho-Jui Hsieh, Xiang-Rui Wang, and Chih-Jen Lin. Lib-linear: A library for large linear classification. *The Journal of Machine Learning Research*, 9:1871–1874, 2008. 4.3.2, 6.2.1
- [65] Alois Ferscha, Clemens Holzmann, and Stefan Oppl. Context awareness for group interaction support. In *Proceedings of the second international workshop on Mobility management & wireless access protocols*, pages 88–97. ACM, 2004. 2.3.1
- [66] Oliver Ferschke, Iris Howley, Gaurav Tomar, Diyi Yang, and Carolyn Rose. Fostering discussion across communication media in massive open online courses. In *CSCL*, 2015. 1, 2.1.2, 2.3.2
- [67] Frank Fischer, Johannes Bruhn, Cornelia Gräsel, and Heinz Mandl. Fostering collaborative knowledge construction with visualization tools. *Learning and Instruction*, 12(2): 213–232, 2002. 3.2.1
- [68] Daniel Fuller and Brian Magerko. Shared mental models in improvisational theatre. In *Proceedings of the 8th ACM conference on Creativity and cognition*, pages 269–278. ACM, 2011. 2.3.1
- [69] D Randy Garrison, Terry Anderson, and Walter Archer. Critical inquiry in a text-based environment: Computer conferencing in higher education. *The internet and higher education*, 2(2):87–105, 1999. 3.2.1
- [70] Elizabeth Avery Gomez, Dezhi Wu, and Katia Passerini. Computer-supported team-based learning: The impact of motivation, enjoyment and team contributions on learning outcomes. *Computers & Education*, 55(1):378–390, 2010. 6.2.1
- [71] Jerry J Gosenpud and John B Washbush. Predicting simulation performance: differences

- between groups and individuals. *Developments in Business Simulation and Experiential Learning*, 18, 1991. 2.3.1
- [72] Sabine Graf and Rahel Bekele. Forming heterogeneous groups for intelligent collaborative learning systems with ant colony optimization. In *Intelligent Tutoring Systems*, pages 217–226. Springer, 2006. 2.3.1
- [73] Sandra Graham and Shari Golan. Motivational influences on cognition: Task involvement, ego involvement, and depth of information processing. *Journal of Educational Psychology*, 83(2):187, 1991. 3.2.1
- [74] Jim Greer, Gordon McCalla, John Cooke, Jason Collins, Vive Kumar, Andrew Bishop, and Julita Vassileva. The intelligent helpdesk: Supporting peer-help in a university course. In *Intelligent tutoring systems*, pages 494–503. Springer, 1998. 2.3.1
- [75] Bonnie Grossen. How should we group to achieve excellence with equity?. 1996. 2.3.1
- [76] Deborah H Gruenfeld, Elizabeth A Mannix, Katherine Y Williams, and Margaret A Neale. Group composition and decision making: How member familiarity and information distribution affect process and performance. *Organizational behavior and human decision processes*, 67(1):1–15, 1996. 6.1
- [77] Gahgene Gweon. *Assessment and support of the idea co-construction process that influences collaboration*. PhD thesis, Carnegie Mellon University, 2012. 1, 2.2, 2.3.2, 6.1
- [78] Gahgene Gweon, Carolyn Rose, Regan Carey, and Zachary Zaiss. Providing support for adaptive scripting in an on-line collaborative learning environment. In *Proceedings of the SIGCHI conference on Human Factors in computing systems*, pages 251–260. ACM, 2006. 2.3.2
- [79] Gahgene Gweon, Mahaveer Jain, John McDonough, Bhiksha Raj, and Carolyn P Rosé. Measuring prevalence of other-oriented transactive contributions using an automated measure of speech style accommodation. *International Journal of Computer-Supported Collaborative Learning*, 8(2):245–265, 2013. 2.2, 3.2.1, 6.2.1
- [80] Linda Harasim, Starr Roxanne Hiltz, Lucio Teles, Murray Turoff, et al. Learning networks, 1995. 2.1.2
- [81] Jeffrey Heer and Michael Bostock. Crowdsourcing graphical perception: using mechanical turk to assess visualization design. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 203–212. ACM, 2010. 3.3.1, 3.3.3
- [82] Khe Foon Hew. Promoting engagement in online courses: What strategies can we learn from three highly rated moocs. *British Journal of Educational Technology*, 2015. 2.1.2
- [83] Iris Howley, Elijah Mayfield, and Carolyn Penstein Rosé. Linguistic analysis methods for studying small groups. *The International Handbook of Collaborative Learning*. 2.3.2
- [84] Mitsuru Ikeda, Shogo Go, and Riichiro Mizoguchi. Opportunistic group formation. In *Artificial Intelligence and Education, Proceedings of AIED*, volume 97, pages 167–174, 1997. 2.3.1
- [85] Albert L Ingram and Lesley G Hathorn. Methods for analyzing collaboration in online. *Online collaborative learning: Theory and practice*, pages 215–241, 2004. 2.3.1, 6.1

- [86] Molly E Ireland and James W Pennebaker. Language style matching in writing: synchrony in essays, correspondence, and poetry. *Journal of personality and social psychology*, 99(3):549, 2010. 5.4.2
- [87] Seiji Isotani, Akiko Inaba, Mitsuru Ikeda, and Riichiro Mizoguchi. An ontology engineering approach to the realization of theory-driven group formation. *International Journal of Computer-Supported Collaborative Learning*, 4(4):445–478, 2009. 2.3.1
- [88] Eugene D Jaffe and Israel D Nebenzahl. Group interaction and business game performance. *Simulation & Gaming*, 21(2):133–146, 1990. 2.3.1
- [89] Namsook Jahng and Mark Bullen. Exploring group forming strategies by examining participation behaviours during whole class discussions. *European Journal of Open, Distance and E-Learning*, 2012. 6.1
- [90] Mahesh Joshi and Carolyn Penstein Rosé. Using transactivity in conversation for summarization of educational dialogue. In *SLaTE*, pages 53–56, 2007. 2.2, 6.2.1, 6.7
- [91] John Keller and Katsuaki Suzuki. Learner motivation and e-learning design: A multinationally validated process. *Journal of Educational Media*, 29(3):229–239, 2004. 4.3.1
- [92] Hanan Khalil and Martin Ebner. Moocs completion rates and possible methods to improve retention-a literature review. In *World Conference on Educational Multimedia, Hypermedia and Telecommunications*, volume 2014, pages 1305–1313, 2014. 2.1.2
- [93] Aniket Kittur and Robert E Kraut. Harnessing the wisdom of crowds in wikipedia: quality through coordination. In *Proceedings of the 2008 ACM conference on Computer supported cooperative work*, pages 37–46. ACM, 2008. 5.4.1
- [94] Aniket Kittur, Ed H Chi, and Bongwon Suh. Crowdsourcing user studies with mechanical turk. In *SIGCHI*, pages 453–456. ACM, 2008. 3.3.1, 3.3.3
- [95] Aniket Kittur, Boris Smus, Susheel Khamkar, and Robert E Kraut. Crowdforge: Crowdsourcing complex work. In *Proceedings of the 24th annual ACM symposium on User interface software and technology*, pages 43–52. ACM, 2011. 9.4.1
- [96] René F Kizilcec and Emily Schneider. Motivation as a lens to understand online learners: Toward data-driven design with the olei scale. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 22(2):6, 2015. 2.2
- [97] Gary G Koch. Intraclass correlation coefficient. *Encyclopedia of statistical sciences*, 1982. 4.3.2
- [98] Daphne Koller, Andrew Ng, Chuong Do, and Zhenghao Chen. Retention and intention in massive open online courses: In depth. *Educause Review*, 48(3):62–63, 2013. 4.1
- [99] Yasmine Kotturi, Chinmay Kulkarni, Michael S Bernstein, and Scott Klemmer. Structure and messaging techniques for online peer learning systems that increase stickiness. In *Proceedings of the Second (2015) ACM Conference on Learning@ Scale*, pages 31–38. ACM, 2015. 2.1.1, 2.2, 8.5
- [100] Steve WJ Kozlowski and Daniel R Ilgen. Enhancing the effectiveness of work groups and teams. *Psychological science in the public interest*, 7(3):77–124, 2006. 2.2

- [101] Robert M Krauss and Susan R Fussell. Perspective-taking in communication: Representations of others' knowledge in reference. *Social Cognition*, 9(1):2–24, 1991. 2.2, 6.1
- [102] Robert E Kraut and Andrew T Fiore. The role of founders in building online groups. In *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing*, pages 722–732. ACM, 2014. 5.4.4, 6
- [103] Chinmay Kulkarni, Julia Cambre, Yasmine Kotturi, Michael S Bernstein, and Scott Klemmer. Talkabout: Making distance matter with small groups in massive classes. In *CSCW*, 2015. 1, 2.1.2
- [104] Rohit Kumar and Carolyn P Rose. Architecture for building conversational agents that support collaborative learning. *Learning Technologies, IEEE Transactions on*, 4(1):21–34, 2011. 2.3.2
- [105] Rohit Kumar and Carolyn P Rosé. Triggering effective social support for online groups. *ACM Transactions on Interactive Intelligent Systems (TiiS)*, 3(4):24, 2014. 2.3.2
- [106] Rohit Kumar, Carolyn Penstein Rosé, Yi-Chia Wang, Mahesh Joshi, and Allen Robinson. Tutorial dialogue as adaptive collaborative learning support. *Frontiers in artificial intelligence and applications*, 158:383, 2007. 1, 2.3.2, 2.4
- [107] Rohit Kumar, Hua Ai, Jack L Beuth, and Carolyn P Rosé. Socially capable conversational tutors can be effective in collaborative learning situations. In *Intelligent tutoring systems*, pages 156–164. Springer, 2010. 2.4
- [108] Theodoros Lappas, Kun Liu, and Evimaria Terzi. Finding a team of experts in social networks. In *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 467–476. ACM, 2009. 2.3.1
- [109] Walter S Lasecki, Young Chol Song, Henry Kautz, and Jeffrey P Bigham. Real-time crowd labeling for deployable activity recognition. In *Proceedings of the 2013 conference on Computer supported cooperative work*, pages 1203–1212. ACM, 2013. 6.3
- [110] Linda Lebie, Jonathan A Rhoades, and Joseph E McGrath. Interaction process in computer-mediated and face-to-face groups. *Computer Supported Cooperative Work (CSCW)*, 4(2-3):127–152, 1995. 5.4.2
- [111] Yi-Chieh Lee, Wen-Chieh Lin, Fu-Yin Cherng, Hao-Chuan Wang, Ching-Ying Sung, and Jung-Tai King. Using time-anchored peer comments to enhance social interaction in online educational videos. In *SIGCHI*, pages 689–698. ACM, 2015. 6.5
- [112] Nan Li, Himanshu Verma, Afroditi Skevi, Guillaume Zufferey, Jan Blom, and Pierre Dillenbourg. Watching moocs together: investigating co-located mooc study groups. *Distance Education*, (ahead-of-print):1–17, 2014. 2.1.3
- [113] Bing Liu, Minqing Hu, and Junsheng Cheng. Opinion observer: analyzing and comparing opinions on the web. In *Proceedings of the 14th international conference on World Wide Web*, pages 342–351. ACM, 2005. 4.3.2
- [114] Edwin A Locke and Gary P Latham. Building a practically useful theory of goal setting and task motivation: A 35-year odyssey. *American psychologist*, 57(9):705, 2002. 4.3.2
- [115] Ioanna Lykourantzou, Angeliki Antoniou, and Yannick Naudet. Matching or crash-



ing? personality-based team formation in crowdsourcing environments. *arXiv preprint arXiv:1501.06313*, 2015. 6.2.1

- [116] Ioanna Lykourantzou, Angeliki Antoniou, Yannick Naudet, and Steven P Dow. Personality matters: Balancing for personality types leads to better outcomes for crowd teams. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*, pages 260–273. ACM, 2016. 2.3.1, 3.3.2
- [117] Ioanna Lykourantzou, Shannon Wang, Robert E Kraut, and Steven P Dow. Team dating: A self-organized team formation strategy for collaborative crowdsourcing. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, pages 1243–1249. ACM, 2016. 2.3.1, 3.3.2
- [118] Sui Mak, Roy Williams, and Jenny Mackness. Blogs and forums as communication and learning tools in a mooc. 2010. 1, 9.4.3
- [119] James R Martin and Peter R White. *The language of evaluation*. Palgrave Macmillan, 2003. 2.2
- [120] Joseph E McGrath. Small group research, that once and future field: An interpretation of the past with an eye to the future. *Group Dynamics: Theory, Research, and Practice*, 1(1):7, 1997. 2.3.2, 5.4.2
- [121] Katelyn YA McKenna, Amie S Green, and Marci EJ Gleason. Relationship formation on the internet: Whats the big attraction? *Journal of social issues*, 58(1):9–31, 2002. 2.3.1
- [122] Larry K Michaelsen and Michael Sweet. Team-based learning. *New directions for teaching and learning*, 2011(128):41–51, 2011. 2.1.4
- [123] Larry K Michaelsen, Arletta Bauman Knight, and L Dee Fink. *Team-based learning: A transformative use of small groups*. Greenwood publishing group, 2002. 6.5
- [124] Colin Milligan, Allison Littlejohn, and Anoush Margaryan. Patterns of engagement in connectivist moocs. *MERLOT Journal of Online Learning and Teaching*, 9(2), 2013. 4.3.1, 4.3.2
- [125] J Moshinskie. How to keep e-learners from escaping. *Performance Improvement Quarterly*, 40(6):30–37, 2001. 4.3.2
- [126] Martin Muehlenbrock. Formation of learning groups by using learner profiles and context information. In *AIED*, pages 507–514, 2005. 2.3.1
- [127] Sean A Munson, Karina Kervin, and Lionel P Robert Jr. Monitoring email to indicate project team performance and mutual attraction. In *Proceedings of the 17th ACM conference on Computer Supported Cooperative Work(CSCW)*, pages 542–549. ACM, 2014. 5.4.2, 5.4.4
- [128] Evelyn Ng and Carl Bereiter. Three levels of goal orientation in learning. *Journal of the Learning Sciences*, 1(3-4):243–271, 1991. 4.3.2
- [129] Kate G Niederhoffer and James W Pennebaker. Linguistic style matching in social interaction. *Journal of Language and Social Psychology*, 21(4):337–360, 2002. 5.4.2
- [130] Bernard A Nijstad and Carsten KW De Dreu. Creativity and group innovation. *Applied*

*Psychology*, 51(3):400–406, 2002. 2.3.1

- [131] Barbara Oakley, Richard M Felder, Rebecca Brent, and Imad Elhajj. Turning student groups into effective teams. *Journal of student centered learning*, 2(1):9–34, 2004. 2.3.1
- [132] Adolfo Obaya. Getting cooperative learning. *Science Education International*, 10(2): 25–27, 1999. 2.3.1
- [133] Asma Ounnas, Hugh Davis, and David Millard. A framework for semantic group formation. In *Advanced Learning Technologies, 2008. ICAIT’08. Eighth IEEE International Conference on*, pages 34–38. IEEE, 2008. 2.3.1
- [134] Pedro Paredes, Alvaro Ortigosa, and Pilar Rodriguez. A method for supporting heterogeneous-group formation through heuristics and visualization. *J. UCS*, 16(19): 2882–2901, 2010. 2.3.1
- [135] James W Pennebaker and Laura A King. Linguistic styles: language use as an individual difference. *Journal of personality and social psychology*, 77(6):1296, 1999. 4.3.2, 4.3.3, 5.4.2
- [136] Normand Roy Ibtihele Bouchoucha Jacques Raynald Jean Talbot Poellhuber, Bruno. and Terry Anderso. The relations between mooc’s participants motivational profiles, engagement profile and persistence. In *MRI Conference, Arlington*, 2013. 4.3.1
- [137] Leo Porter, Cynthia Bailey Lee, and Beth Simon. Halving fail rates using peer instruction: a study of four computer science courses. In *Proceeding of the 44th ACM technical symposium on Computer science education*, pages 177–182. ACM, 2013. 2.1.4
- [138] Arti Ramesh, Dan Goldwasser, Bert Huang, Hal Daume III, and Lise Getoor. Modeling learner engagement in moocs using probabilistic soft logic. In *NIPS Workshop on Data Driven Education*, 2013. 4.1, 4.3.2, 4.4.3
- [139] Keith Richards. *Language and professional identity: Aspects of collaborative interaction*. Palgrave Macmillan, 2006. 2.2
- [140] Elena Rocco. Trust breaks down in electronic contexts but can be repaired by some initial face-to-face contact. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 496–502. ACM Press/Addison-Wesley Publishing Co., 1998. 2.1.2
- [141] Farnaz Ronaghi, Amin Saberi, and Anne Trumbore. Novoed, a social learning environment. *Massive Open Online Courses: The MOOC Revolution*, page 96, 2014. 6.1
- [142] Carolyn Rosé, Ryan Carlson, Diyi Yang, Miaomiao Wen, Lauren Resnick, Pam Goldman, and Jennifer Sheerer. Social factors that contribute to attrition in moocs. In *ACM Learning at Scale*, 2014. 3.2.2
- [143] Carolyn Rosé, Yi-Chia Wang, Yue Cui, Jaime Arguello, Karsten Stegmann, Armin Weinberger, and Frank Fischer. 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, 2008. 2.2
- [144] Michael F Schober. Spatial perspective-taking in conversation. *Cognition*, 47(1):1–24,

1993. 2.2, 6.1

- [145] Michael F Schober and Susan E Brennan. Processes of interactive spoken discourse: The role of the partner. *Handbook of discourse processes*, pages 123–164, 2003. 2.2, 6.1
- [146] Donald A Schön. *The reflective practitioner: How professionals think in action*, volume 5126. Basic books, 1983. 2.3.1
- [147] Daniel L Schwartz. The productive agency that drives collaborative learning. *Collaborative learning: Cognitive and computational approaches*, pages 197–218, 1999. 3.2.1, 6.1
- [148] Burr Settles and Steven Dow. Let’s get together: the formation and success of online creative collaborations. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 2009–2018. ACM, 2013. 2.3.1
- [149] J Bryan Sexton and Robert L Helmreich. Analyzing cockpit communications: the links between language, performance, error, and workload. *Journal of Human Performance in Extreme Environments*, 5(1):6, 2000. 5.4.2
- [150] George Siemens. Connectivism: A learning theory for the digital age. *International journal of instructional technology and distance learning*, 2(1):3–10, 2005. 2.1.4
- [151] Judith D Singer and John B Willett. *Applied longitudinal data analysis: Modeling change and event occurrence*. Oxford university press, 2003. 5.2.1
- [152] Glenn Gordon Smith, Chris Sorensen, Andrew Gump, Allen J Heindel, Mieke Caris, and Christopher D Martinez. Overcoming student resistance to group work: Online versus face-to-face. *The Internet and Higher Education*, 14(2):121–128, 2011. 2.2
- [153] Michelle K Smith, William B Wood, Wendy K Adams, Carl Wieman, Jennifer K Knight, Nancy Guild, and Tin Tin Su. Why peer discussion improves student performance on in-class concept questions. *Science*, 323(5910):122–124, 2009. 2.1.4
- [154] Rion Snow, Brendan O’Connor, Daniel Jurafsky, and Andrew Y Ng. Cheap and fast—but is it good?: evaluating non-expert annotations for natural language tasks. In *Proceedings of the conference on empirical methods in natural language processing*, pages 254–263. Association for Computational Linguistics, 2008. 4.3.2
- [155] Leen-Kiat Soh, Nobel Khandaker, Xuliu Liu, and Hong Jiang. A computer-supported cooperative learning system with multiagent intelligence. In *Proceedings of the fifth international joint conference on Autonomous agents and multiagent systems*, pages 1556–1563. ACM, 2006. 2.3.1
- [156] Amy Soller, Alejandra Martínez, Patrick Jermann, and Martin Muehlenbrock. From mirroring to guiding: A review of state of the art technology for supporting collaborative learning. *International Journal of Artificial Intelligence in Education (IJAIED)*, 15:261–290, 2005. 2.3.2
- [157] Gerry Stahl. *Group cognition: Computer support for building collaborative knowledge*. Mit Press Cambridge, MA, 2006. 2.2, 2.3.2
- [158] Gerry Stahl. Learning across levels. *International Journal of Computer-Supported Collaborative Learning*, 8(1):1–12, 2013. 2.1.3

- [159] Thomas Staubitz, Jan Renz, Christian Willems, and Christoph Meinel. Supporting social interaction and collaboration on an xmooc platform. *Proc. EDULEARN14*, pages 6667–6677, 2014. 2.1.1, 2.2
- [160] Ivan D Steiner. Group process and productivity (social psychological monograph). 2007. 2.2
- [161] Sue Stoney and Ron Oliver. Can higher order thinking and cognitive engagement be enhanced with multimedia. *Interactive Multimedia Electronic Journal of Computer-Enhanced Learning*, 1(2), 1999. 4.3.3
- [162] Jan-Willem Strijbos. *THE EFFECT OF ROLES ON COMPUTER-SUPPORTED COLLABORATIVE LEARNING*. PhD thesis, Open Universiteit Nederland, 2004. 2.1.4
- [163] James T Strong and Rolph E Anderson. Free-riding in group projects: Control mechanisms and preliminary data. *Journal of Marketing Education*, 12(2):61–67, 1990. 2.3.1
- [164] Pei-Chen Sun, Hsing Kenny Cheng, Tung-Chin Lin, and Feng-Sheng Wang. A design to promote group learning in e-learning: Experiences from the field. *Computers & Education*, 50(3):661–677, 2008. 6
- [165] Daniel D Suthers. Technology affordances for intersubjective meaning making: A research agenda for cscl. *International Journal of Computer-Supported Collaborative Learning*, 1(3):315–337, 2006. 3.2.1, 6.1
- [166] Robert I Sutton and Andrew Hargadon. Brainstorming groups in context: Effectiveness in a product design firm. *Administrative Science Quarterly*, pages 685–718, 1996. 2.3.1
- [167] Karen Swan, Jia Shen, and Starr Roxanne Hiltz. Assessment and collaboration in online learning. *Journal of Asynchronous Learning Networks*, 10(1):45–62, 2006. 6.5
- [168] Luis Talavera and Elena Gaudio. Mining student data to characterize similar behavior groups in unstructured collaboration spaces. In *Workshop on artificial intelligence in CSCL. 16th European conference on artificial intelligence*, pages 17–23, 2004. 2.3.2
- [169] Yla R Tausczik and James W Pennebaker. Improving teamwork using real-time language feedback. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 459–468. ACM, 2013. 2.3.2, 5.4.2
- [170] Yla R Tausczik, Laura A Dabbish, and Robert E Kraut. Building loyalty to online communities through bond and identity-based attachment to sub-groups. In *Proceedings of the 17th ACM conference on Computer supported cooperative work and social computing*, pages 146–157. ACM, 2014. 2.1.2
- [171] Stephanie D Teasley. Talking about reasoning: How important is the peer in peer collaboration? In *Discourse, Tools and Reasoning*, pages 361–384. Springer, 1997. 3.2.1
- [172] Stephanie D Teasley, Frank Fischer, Armin Weinberger, Karsten Stegmann, Pierre Dillenbourg, Manu Kapur, and Michelene Chi. Cognitive convergence in collaborative learning. In *International conference for the learning sciences*, pages 360–367, 2008. 2.2, 6.1, 6.7
- [173] Lydia T Tien, Vicki Roth, and JA Kampmeier. Implementation of a peer-led team learning instructional approach in an undergraduate organic chemistry course. *Journal of research in science teaching*, 39(7):606–632, 2002. 9.4.3

- [174] Vincent Tinto. *Leaving college: Rethinking the causes and cures of student attrition*. ERIC, 1987. 2.1.2
- [175] Gaurav Tomar, Xu Wang Sreecharan Sankaranarayanan, and Carolyn Penstein Ros. Coordinating collaborative chat in massive open online courses. In *CSCS*, 2016. 2.1.3
- [176] Gaurav Singh Tomar, Sreecharan Sankaranarayanan, and Carolyn Penstein Rosé. Intelligent conversational agents as facilitators and coordinators for group work in distributed learning environments (moocs). In *2016 AAAI Spring Symposium Series*, 2016. 2.1.2
- [177] W Ben Towne and James D Herbsleb. Design considerations for online deliberation systems. *Journal of Information Technology & Politics*, 9(1):97–115, 2012. 10.3.3
- [178] Sherry Turkle. *Alone together: Why we expect more from technology and less from each other*. Basic books, 2012. 2.1.1
- [179] Peter D Turney, Yair Neuman, Dan Assaf, and Yohai Cohen. Literal and metaphorical sense identification through concrete and abstract context. In *Proceedings of the 2011 Conference on the Empirical Methods in Natural Language Processing*, pages 680–690, 2011. 4.3.3
- [180] Judith Donath Karrie Karahalios Fernanda Viegas, J Donath, and K Karahalios. Visualizing conversation, 1999. 2.3.2
- [181] Ana Cláudia Vieira, Lamartine Teixeira, Aline Timóteo, Patrícia Tedesco, and Flávia Barros. Analyzing on-line collaborative dialogues: The oxentchê-chat. In *Intelligent Tutoring Systems*, pages 315–324. Springer, 2004. 2.3.2
- [182] Luis Von Ahn and Laura Dabbish. Labeling images with a computer game. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 319–326. ACM, 2004. 3.3.2
- [183] Lev S Vygotsky. *Mind in society: The development of higher psychological processes*. Harvard university press, 1980. 2.1.4
- [184] Erin Walker. *Automated adaptive support for peer tutoring*. PhD thesis, Carnegie Mellon University, 2010. 2.3.2
- [185] Xinyu Wang, Zhou Zhao, and Wilfred Ng. Ustf: A unified system of team formation. *IEEE Transactions on Big Data*, 2(1):70–84, 2016. 2.3.1
- [186] Xu Wang, Diyi Yang, Miaomiao Wen, Kenneth Koedinger, and Carolyn P Rosé. Investigating how students cognitive behavior in mooc discussion forums affect learning gains. *The 8th International Conference on Educational Data Mining*, 2015. 4
- [187] Yi-Chia Wang, Robert Kraut, and John M Levine. To stay or leave?: the relationship of emotional and informational support to commitment in online health support groups. In *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work*, pages 833–842. ACM, 2012. 3.2.2
- [188] Noreen M Webb. The teacher’s role in promoting collaborative dialogue in the classroom. *British Journal of Educational Psychology*, 79(1):1–28, 2009. 2.3
- [189] Carine G Webber and Maria de Fátima Webber do Prado Lima. Evaluating automatic

- group formation mechanisms to promote collaborative learning—a case study. *International Journal of Learning Technology*, 7(3):261–276, 2012. 2.3.1
- [190] Armin Weinberger and Frank Fischer. A framework to analyze argumentative knowledge construction in computer-supported collaborative learning. *Computers & education*, 46(1):71–95, 2006. 1, 2.3.2, 3.2.1, 6.1
- [191] Armin Weinberger, Karsten Stegmann, Frank Fischer, and Heinz Mandl. Scripting argumentative knowledge construction in computer-supported learning environments. In *Scripting computer-supported collaborative learning*, pages 191–211. Springer, 2007. 2.3.2
- [192] Miaomiao Wen and Carolyn Penstein Rosé. Identifying latent study habits by mining learner behavior patterns in massive open online courses. In *Proceedings of the 23rd ACM International Conference on Conference on Information and Knowledge Management*, pages 1983–1986. ACM, 2014. 1
- [193] Miaomiao Wen, Diyi Yang, and Carolyn Rosé. Linguistic reflections of student engagement in massive open online courses. In *International AAAI Conference on Weblogs and Social Media*, 2014. 1, 5.2.1
- [194] Miaomiao Wen, Diyi Yang, and Carolyn Penstein Rosé. Sentiment analysis in mooc discussion forums: What does it tell us. *Proceedings of Educational Data Mining*, 2014. 1
- [195] Miaomiao Wen, Diyi Yang, and Carolyn Penstein Rosé. Virtual teams in massive open online courses. *Proceedings of Artificial Intelligence in Education*, 2015. 1
- [196] Martin Wessner and Hans-Rüdiger Pfister. Group formation in computer-supported collaborative learning. In *Proceedings of the 2001 international ACM SIGGROUP conference on supporting group work*, pages 24–31. ACM, 2001. 2.3.1
- [197] Ian H Witten and Eibe Frank. *Data Mining: Practical machine learning tools and techniques*. Morgan Kaufmann, 2005. 4.3.2
- [198] Joseph Wolfe and Thomas M Box. Team cohesion effects on business game performance. *Developments in Business Simulation and Experiential Learning*, 14, 1987. 2.3.1
- [199] Anita Williams Woolley, Christopher F Chabris, Alex Pentland, Nada Hashmi, and Thomas W Malone. Evidence for a collective intelligence factor in the performance of human groups. *science*, 330(6004):686–688, 2010. 6.4.2
- [200] Diyi Yang, Tanmay Sinha, David Adamson, and Carolyn Penstein Rosé. Turn on, tune in, drop out: Anticipating student dropouts in massive open online courses. In *Workshop on Data Driven Education, Advances in Neural Information Processing Systems 2013*, 2013. 3.2.2, 4.4.1
- [201] Diyi Yang, Miaomiao Wen, and Carolyn Rose. Peer influence on attrition in massive open online courses. In *Proceedings of Educational Data Mining*, 2014. 5.3.2
- [202] Robert K Yin. The case study method as a tool for doing evaluation. *Current Sociology*, 40(1):121–137, 1992. 3.4
- [203] Robert K Yin. *Applications of case study research*. Sage, 2011. 3.4

- [204] Zehui Zhan, Patrick SW Fong, Hu Mei, and Ting Liang. Effects of gender grouping on students group performance, individual achievements and attitudes in computer-supported collaborative learning. *Computers in Human Behavior*, 48:587–596, 2015. 2.3.1
- [205] Jianwei Zhang, Marlene Scardamalia, Richard Reeve, and Richard Messina. Designs for collective cognitive responsibility in knowledge-building communities. *The Journal of the Learning Sciences*, 18(1):7–44, 2009. 2.3.1
- [206] Zhilin Zheng, Tim Vogelsang, and Niels Pinkwart. The impact of small learning group composition on student engagement and success in a mooc. *Proceedings of Educational Data Mining*, 7, 2014. 1, 2.3.1, 6
- [207] Zhilin Zheng, Tim Vogelsang, and Niels Pinkwart. The impact of small learning group composition on student engagement and success in a mooc. *Proceedings of Educational Data Mining*, 2014. 8.5
- [208] Erping Zhu. Interaction and cognitive engagement: An analysis of four asynchronous online discussions. *Instructional Science*, 34(6):451–480, 2006. 3.2.1, 4.3.3
- [209] Haiyi Zhu, Steven P Dow, Robert E Kraut, and Aniket Kittur. Reviewing versus doing: Learning and performance in crowd assessment. In *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing*, pages 1445–1455. ACM, 2014. 6.2.1
- [210] Barry J Zimmerman. A social cognitive view of self-regulated academic learning. *Journal of educational psychology*, 81(3):329, 1989. 2.1.1
- [211] Jörg Zumbach and Peter Reimann. Influence of feedback on distributed problem based learning. In *Designing for change in networked learning environments*, pages 219–228. Springer, 2003. 2.3.1