Watching MOOCs Together: Investigating Co-located MOOC Study Groups

Nan Li, Himanshu Verma, Afroditi Skevi, Guillaume Zufferey, Jan Blom, Pierre Dillenbourg

Computer-Human Interaction in Learning and Instruction Lab (CHILI), École Polytechnique Fédérale de Lausanne, Switzerland

EPFL-CHILI, RLC D1 740 (Rolex Learning Center), Station 20, CH-1015 Lausanne, Vaud, Switzerland

nan.li@epfl.ch, himanshu.verma@epfl.ch, afroditi.skevi@epfl.ch, guillaume.zufferey@epfl.ch, jan.blom@google.com, pierre.dillenbourg@epfl.ch

Watching MOOCs Together:

Investigating Co-located MOOC Study Groups

Current MOOCs emphasize offering individualistic learning experiences, and most MOOC research is centered on this context. Recent research also suggests that MOOC students may prefer to study in groups. Social facilitation within the study groups may render the learning of difficult concepts a pleasing experience. We report a longitudinal study that investigates how co-located study groups watch and study MOOC videos together. The study was conducted with on-campus flipped classroom students. Our subjects reported an overall high satisfaction with the study group style. The research reveals that students like to stay synchronized in the group while watching MOOC videos. However, they have to find a balance between synchronization, video interaction and the amount of conversation. Watching MOOCs on a shared display addresses the need of synchronicity, and the distribution of control can increase the video interactivities.

Keywords: MOOC, collaborative video watching, study groups

Introduction

MOOCs' large audience base is challenging the pedagogical practices. One prominent problem is the elimination of the traditional interaction between instructors and learners, so that learning feedback cannot be obtained directly from the instructors (Kop, Fournier & Mak, 2011). Therefore, learners often seek support from their peers. The benefits of such collaboration have been established by several learning theories, such as social constructivism (Vygotsky, 1978) and connectivism (Siemens, 2005). Through collaboration, the ideal outcome is to achieve a situation in which "*the learner is the teacher is the learner*" (Siemens, 2006). Most popular MOOC platforms provide online forums to foster collaboration among learners. With the forums, students create temporary social interactions that are asynchronous and voluntary - timely support is not guaranteed. Furthermore, learner diversity makes it harder to maintain the activeness and quality of the loose collaborations within a forum.

Collaborative Learning in Study Groups

Research has found that the more open an online course is, the more the learners seek to engage in groups, as opposed to an open network (Mackness, Mak & Williams, 2010). This suggests that MOOC learners may be inclined towards learning in groups. Apart from students' subjective willingness, many researchers have identified (over decades) the benefits of such collaborations in terms of learning effectiveness. Johnson & Johnson (1989) found that learning is most effective when students work collaboratively, share and compare individual ideas and work on a common solution. Furthermore, students develop critical skills and construct common knowledge, when learning in groups (Brindley, Blaschke & Walti, 2009).

While the power of group learning has been demonstrated, how can we realize it in the context of MOOCs? One intuitive solution is to bring group engagement online. Researchers have investigated online group learning for years. Learners reported to experience more frustrations in online groups than in the face-to-face counterpart (Smith et al., 2011), which may relate to the difference in study goals, imbalanced

participation and the quality of individual contribution (Capdeferro & Romero, 2012). NovoED[1] (a MOOC initiative) created a social incentive system to tackle these challenges: Small group collaborations are enforced and implemented via Google Doc[2] and Hangouts[3]. Individual performance in a group is peer-rated so as to encourage participation and contribution. NovoED currently focuses on entrepreneurship courses consisting of designated group projects. In comparison, courses in technology and mathematics do not always have suitable group tasks. As MOOC learning is often centred on lecture videos, a more general group-based learning pedagogical approach could leverage this activity.

Collaborative Video Watching

Over the last 40 years, extensive research has been conducted in the field of collaborative video watching. In the late 70s, Gibbons, Kincheloe & Down (1977) coined the term "Tutored Video Instruction (TVI)" to denote the scenario where remote students watch video lectures in small groups with a tutor. Both students and tutors were able to pause the video and initiate discussions when problems and questions arose. The result showed that in terms of average grade obtained by the students, TVI students outperformed students who watched live video lectures in the classroom and those who watched offline video lectures. More surprisingly, TVI students also outperformed campus students who attended the lecture in the classroom. While tutors were present in this research study, the scenario where group students watched lectures without a tutor was also reported as advantageous in another research study (Stone, 1990). Sun (Sipusic et al., 1999) and Microsoft researchers (Smith, Sipusic & Pannoni, 1999) also demonstrated the advantages of synchronous collaborative video-watching of lectures, by replicating TVI in remote settings (Distributed-TVI or DTVI). A subsequent research on DTVI without tutors again confirmed similar conclusions (Cadiz et al., 2000). A new term collaborative video viewing (CVV) was coined later to represent the scenario where TVI is conducted without tutors. Similar to DTVI as compared with TVI, Distributed-CVV (DCVV) is used in the distributed condition. In addition, Cadiz et al. (2000) compared CVV with DCVV on learning and interaction behavior. Their results exhibit that the collocated groups were significantly more comfortable with pausing videos and discussed for longer duration both in total and per pause, than the distributed groups. Furthermore, another follow-up research by Weisz et al. (2007) showed that discussing while video watching was perceived to be an engaging and enriching social experience by the participants and was not considered distracting.

Early research on TVI required the presence of a tutor, which is not a realistic solution for MOOCs. DCVV can be potentially feasible, but current MOOCs do not technically offer synchronous online collaborative video watching experiences. In contrast, CVV can be achieved for students that are close to each other.

As MOOCs have reached large scale of learners, geographical student clusters may emerge. This phenomenon is seen in MeetUp[4], which is a communication hub for learners in the same location. We have observed that while local meetings are actually being organized spontaneously, they are mostly unstructured and casual. MeetUp in its current form does not provide suggestions on how to study together.

---

[1] NovoED: https://novoed.com/
[2] Google Docs: https://docs.google.com/
[3] Google Hangouts: http://www.google.com/hangouts/
[4] MeetUp: http://www.meetup.com/

In addition to MeetUp, universities naturally gather learners. It is highly probable that a student is following the same MOOC course as other fellow students/friends. Furthermore, many universities are embedding MOOCs into their course schedule (Martin, 2012) with the flipped-classroom model (Tucker, 2012). This model offers more group learning opportunities for university students, who are used to spontaneously studying in groups. Such emergent study groups are also shown to be effective in achieving better outcomes than individual learners (Tang, 1993), in terms of grades.

Given the possibility for students to study together in the context of MOOCs, we are motivated to replicate the CVV approach in MOOC learning. We conducted an experiment with spontaneously formed study groups of on-campus flipped-classroom students. In this article, we report our findings on how the students watched lecture videos and collaborated in groups over several weeks. On-campus students are representatives of collocated MOOC learners. We believe the investigation of this subpopulation of MOOC learners has pedagogical implications for flipped-classroom teaching, the organization of MeetUps as well as distance educational programs for developing countries where digital infrastructure is limited. It also has technological implications towards the design of video player features customized for watching MOOC lectures in a distributed group.

Research Questions

We designed a user study to investigate the potential of watching MOOC video lectures as well as solving quizzes collaboratively, within a collocated study-group setting. It is expected that study groups might adopt different styles of watching MOOC video lectures. For example, some groups might decide to watch video lectures independently followed by discussions after the end of each video. Other groups might prefer watching video lectures synchronously by projecting the video on a screen. Considering these varied video-watching styles, we defined different video-watching conditions in our study by manipulating across two dimensions: *display* (centralized or distributed) and *video control* (shared or distributed) as explained in the next sections.

In order to investigate the underlying video-watching dynamics of the study groups, we aim to answer the following research questions, clustered into the three aspects of group work. The research questions were chosen to be thematically broad, as we intended to explore various factors involved in the study, and to get a more realistic picture of underlying study-group behaviour.

a) **Synchronicity:** The varied video-watching styles adopted by the study groups reflect on the MOOC video watching preferences. One such preference worth investigating is the synchronous watching of video lectures. *Do study groups tend to watch videos asynchronously (independent watching within the time frame when the students meet to study), or is synchronicity a desirable attribute of group video watching?*

b) **Discussions and argumentations:** In collocated study-group settings, discussions among the group members are expected, especially while watching the difficult parts of the video lectures. Concerning discussions, we aim to examine: *When do the group members usually discuss? Are there discussions while watching a video, or after finishing a video lecture? Further, is there an influence of the video-watching configuration on the discussion patterns of the group members?*

c) **Navigation patterns:** Finally, we anticipate that some groups might interact more frequently with the video lectures by pausing and seeking, than others. One can think of several factors responsible for this behaviour, such as group composition or course difficulty. However, the influence of different video-watching conditions on the video interactions is more interesting to investigate, as this might provide some critical insights into the identification of an effective video watching condition. Thus, our question is: *Do the different video-watching styles lead to difference in the amount of interactivity with the video lectures?*

Experimental Methodology

We conducted an exploratory longitudinal study lasting for 5 weeks in the spring of 2013. During these 5 weeks the study groups were asked to watch video lectures collaboratively and solve quizzes, subjected to three video-watching conditions.

Participants

Fifty-Four (54) engineering students (11 females, 43 males) were recruited from École Polytechnique Fédérale de Lausanne (EPFL) to participate in our study. Twelve (12) groups of 4-5 participants were spontaneously formed. These 12 groups attended the weekly lectures from one of two MOOCs, which were offered in a flipped classroom format; namely *Numerical Analysis* (NAS, in French) and *Digital Signal Processing* (DSP, in English). The participants were recruited from amongst the pool of students who had enrolled in these courses on campus, and were asked to form their own study-groups (except one group, which was formed by the experimenters). Both the courses were mathematical in nature, and the students could meet with the instructors weekly in a classroom. The students enrolled in the NAS course were first-year bachelor students, whereas the students enrolled in the DSP course consisted of first- and second-year master students. Also, the gender ratio of the participants in the study was unbalanced, which is a drawback of any authentic learning context. More details about group distribution are given in Table 1.

Most of the students who participated in our study had no previous experience of learning with MOOCs. In addition, all except one student reported having participated in a study-group in the past. Further, all group members with an exception of one group from DSP course (which was formed by us), reported that they were well acquainted with their respective group members. Finally, each participant was rewarded with a compensation amount of 150 Swiss Francs and a printed textbook of the course, after the completion of the experiment.

Table 1

Distribution of participants across the two courses.

Course	Total Groups	Females	Males
NAS	9	11	29
DSP	3	0	14

Experimental Conditions

We decided to have three experimental conditions, based on the groups' varied video-watching styles. The choice of these three conditions was done based on manipulations across two dimensions: *display* and *video controller* as shown in Table 2. The *display*

dimension refers to the screen over which the group members watch the video lectures. The *video controller* dimension refers to the controller that enables group members to interact with the video by pausing, seeking and browsing between videos. Both these dimensions can be either *centralized* or *distributed*.

Table 2

The distribution of experimental conditions across the two dimensions.

| | Display | |
Video Controller	Centralized	Distributed
Centralized	CC	-
Distributed	DD	DC

Varying across the aforementioned dimensions, we define three experimental conditions (see Figure 1) as follows:

- **Centralized video control and centralized display (CC):** In this condition, group members sat around a table and watched video lectures on a single wall-mounted projected display, connected to a single tablet computer (multi-touch enabled). The group members could control the video by using the provided video-controller in the video player application. This condition is similar to the CVV condition from the user study of Cadiz et al. (2000).
- **Distributed video control and centralized display (DC):** In this condition, group members sat around a table and watched videos on a single wall-mounted projected display, connected to a PC. In addition, each learner was equipped with a mouse of her own, which could be used to interact with the video player. This condition can be considered to be similar to a single television with multiple remote controllers, where each viewer gets her own remote control.
- **Distributed video control and distributed display (DD):** In this condition, group members sat around the table and watched the video lectures on a separate tablet computer (multi-touch enabled) at their own pace. Each group member could only control her own video. Further, the participants in this condition used headphones while watching video lectures.

(a) CC Condition: Group watch video on the centralized display with a single video control

(b) DC Condition: Centralized display but each group member has their own video control

(c) DD Condition: Every group member watches the video on a different tablet computer

Figure 1. The three experimental conditions

We envisioned that the DC condition would be an intermediate condition imposing synchronous video watching while encouraging group members to interact with the video by providing individual controls. This condition could be a good supplement to the study-group way of watching videos, as providing each group

member with a video controller of their own might encourage them to pause and discuss frequently. We chose to omit the CD (centralized video control and distributed display) condition, as this condition does not correspond to a realistic study-group practice.

Methodology

The 12 groups were evenly distributed across the three conditions, such that each condition had 4 groups (3 groups of NAS, 1 group of DSP). The study lasted for a period of 5 weeks out of 7 for the two MOOC courses. Each study-group was asked to watch MOOC videos and solve quizzes collaboratively. The participants were instructed well in advance not to watch the MOOC videos before coming for the weekly experiment session.

We used one-factor (video watching condition), between-group design for our study. Participants were allowed to discuss anytime while watching videos and were motivated to do so. The experiment sessions were not time-bound and participants were asked to watch videos at their own pace. However, an upper limit of 3 hours was set for a single experiment session. The video length of each week's lectures was less than the time limit set for each session (see Table 3). After each session, participants completed a post-experiment questionnaire recording their perceptions of the session.

We observed the study-groups in their natural behaviour and the participants were asked to choose their style freely. However, as our study was exploratory in nature and focused on the collaborative video watching, we decided not to measure any learning gain of the participants.

Table 3

The total video length (in minutes) of each weekly session

Course	Week 1	Week 2	Week 3	Week 4	Week 5	Mean (SD)
NAS	39.36	56.90	59.26	48.28	59.46	52.65 (8.72)
DSP	58.35	111.76	94.93	107.21	135.21	101.49 (28.19)

Data Collected

We recorded all the interactions of the participants with the video player (pauses, seek, etc.) in the system log files. In addition, as each participant had her own video controller (mouse) in DC condition, we also recorded individual actions in the log files. The participants completed a pre-experiment questionnaire, recording basic demographics and personality information (such as participant's perception of herself about extroversion, self-discipline, etc.) as well as their previous experiences with MOOCs and study groups. The participants also filled in a post-experiment questionnaire after each weekly session, stating their perceptions about the video watching session, satisfaction and quality of their discussions. Finally, we video recorded all experimental sessions and also conducted semi-structured interviews with the participants at the end of the user study.

Results and Analysis

Participants' Subjective Perceptions

The experiment participants reported on a high level of satisfaction with the study-group way of learning with MOOCs, in the post-experiment questionnaires as well as in the semi-structured interviews. The discussions were perceived to be of higher quality and very beneficial on a 5-point Likert-scale (see Table 4), even in the DD condition where the group members were watching the videos independently. We observed that in the DD condition, study group members were waiting for others to finish their respective videos and discussions followed after each video. This shows that being in the study group influenced the individual video-watchers to synchronize their studying practice while giving a significant priority to the discussions after each video. Also, we did not find any significant statistical difference across the 3 conditions regarding the difference in perceived discussion quality.

Table 4

The perceived quality of discussions during MOOC video watching. No significant difference was found in perceived quality across conditions.

Perceived Quality of Discussions	Mean (SD)
All Participants	4.12 (0.81)
CC Condition	4.17 (0.79)
DC Condition	4.14 (0.87)
DD Condition	4.03 (0.76)

During the semi-structured interviews, participants reported that the power to pause the MOOC lecture in situations where a learner needs explanations about a specific concept, followed by discussions with peers was similar in experience to "*…it's like pausing a professor in the classroom.*". The participants also had positive opinions concerning their motivations for learning with MOOCs in a study group as they can validate their understanding of the topic with their friends instantly. This was reported by one of the participant as "*If you are alone, maybe you doubt about your results and understanding and you validate the solutions to problems two or three times, to be sure. But here [within study group] you can compare with your friends. This is much better!*". Moreover, studying individually with MOOCs was perceived to be less effective and less motivating than learning within a study group. This was reported as: "*The study group is much better than studying alone. More motivating … asking questions and getting answers, so that we are able to understand better. Alone, it would be more difficult and less interesting*". Study participants reported that as many courses are difficult and time-consuming, the study group renders their learning experience more pleasing. Finally, participants reported of increased activeness and attentiveness while watching video-lectures in a group as: "*… during lectures [classroom] sometimes you kind of drift away, while here [study group] because of the interaction with the group, you are obliged to be more active*".

Video Navigation Patterns

In order to gain perceptual knowledge of how participants interacted with video-lectures, we visualize students' navigation within videos with video navigation plots. Four sample plots, each illustrating the navigation patterns for a single group from different conditions, are shown in Figure 2.
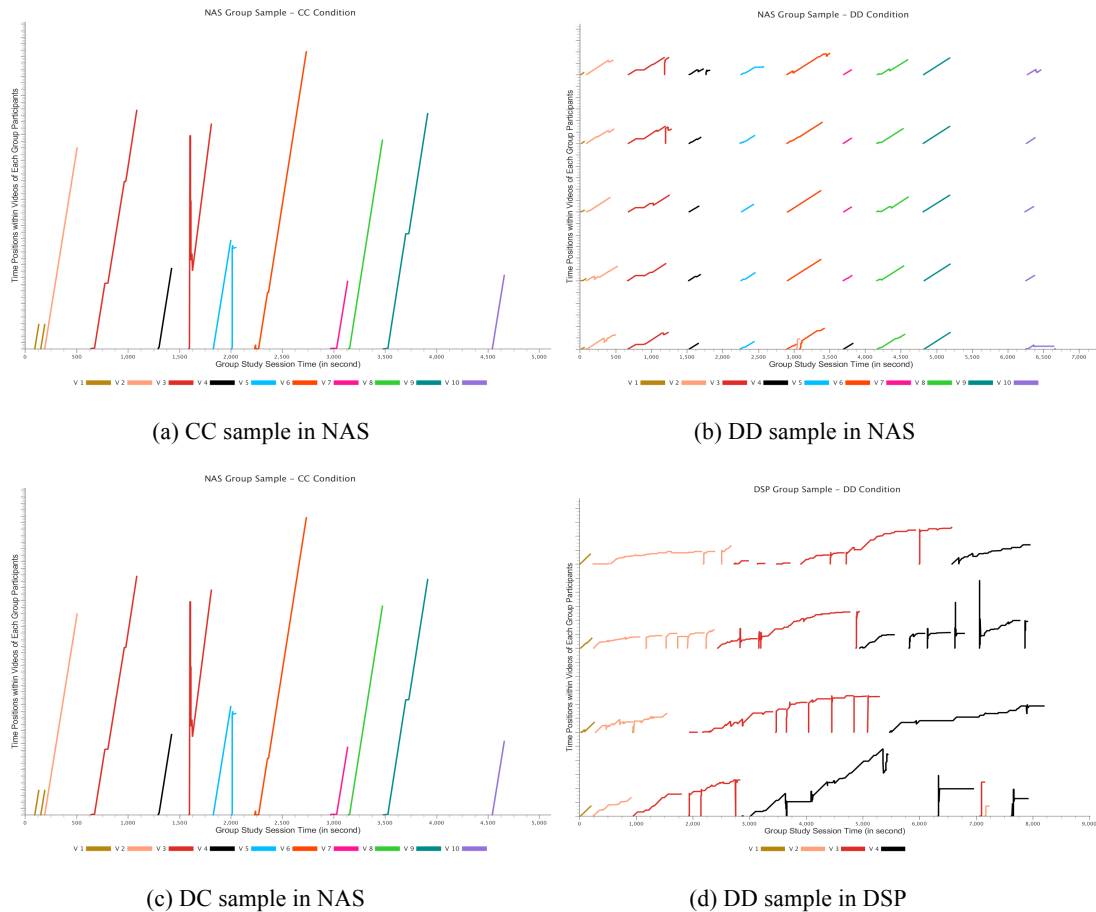

(a) CC sample in NAS

(b) DD sample in NAS

(c) DC sample in NAS

(d) DD sample in DSP

Figure 2. Sample video navigation patterns of the study groups for a weekly study session. The horizontal axis represents the session time (in seconds). The vertical axis corresponds to the video-time position of a student/group (in seconds). Each video is associated with a specific color.

The horizontal axis of each plot represents the timeline of a study group session, and the vertical axis denotes the timeline of the videos. Both timelines are measured in seconds. Groups watched multiple videos each week, and each color in the plot represents a different video. Figure 2(b) and 2(d) are the examples of DD condition, where the patterns for multiple students are shown in parallel. A straight line-segment with a positive slope indicates that the corresponding video was played without any interruption; a horizontal line-segment is a sign of a pause; jitters depict fast forwards/backwards within the video and the gaps between two continuous series are the time periods when students were discussing about the problems or doing quizzes after finishing a video (no active video at that moment). Participants did not ask for breaks, so the plots represent a complete picture of the group activity during a session.

Comparing the patterns in DD condition for the NAS course (Figure 2(b)) with the other two conditions for the same course (Figure 2(a) and 2(c)), we do not see striking differences. It is the DD group for the DSP course (Figure 2(d)) that stands out from the four samples. Students in this DSP group interacted with video-lectures a lot more, and were watching different video content at the same time, than those from the NAS course. Unlike NAS groups, other DSP groups exhibit differing patterns as well. Moreover, the instructors of the DSP course did not oblige campus students to solve MOOC quizzes, and the content of the course was more advanced than the introductory NAS course. These factors constitute a strong course-dependent effect that would break the homogeneity of the analysis. Considering that only 3 DSP groups were recruited, we have less statistical power to prove a finding. Therefore, the analyses in the next sections will be solely based on NAS groups.

Video Interactivity

Participants' interaction with video-lectures is presented as video navigation plots in Figure 2. Video interactions consist of different kinds of actions (pause, seek-forward and seek-backward), each contributing to the total video-watching time. We define **time-spent-on-video index (TSOVI)** to gauge the level of interactivity. TSOVI refers to the ratio between the total time spent on watching videos in a week and the total length of video contents that are watched (not necessarily full videos). Possible values are theoretically any numbers that are above 1.0. Both pausing and rewinding videos increase the value: an index of 1.0 indicates that all the videos were played exactly once without being paused or re-watched, otherwise it exhibits that students had spent additional time on the videos. While computing the indices, all the videos for a week were taken into account as a whole, and non-watched video content was not considered.

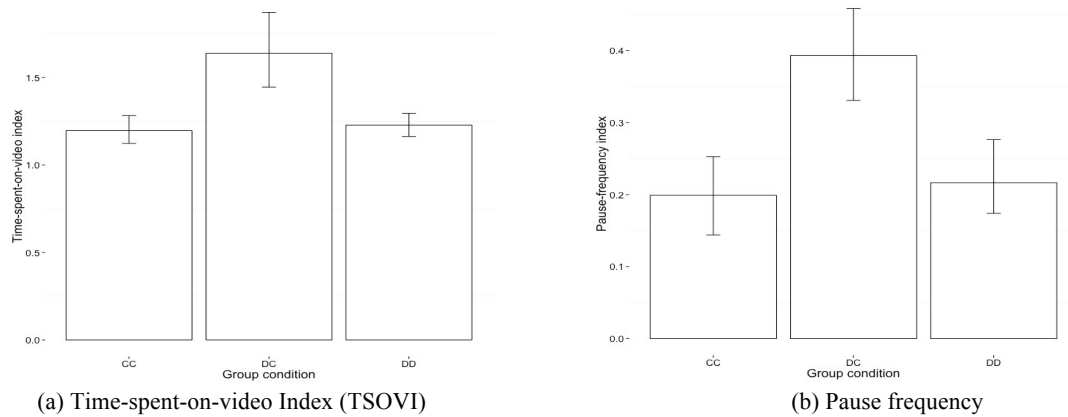(a) Time-spent-on-video Index (TSOVI)   (b) Pause frequency

Figure 3. Bar plots with means and confidence intervals for time-spent-on-video index and pause frequency under different conditions

Descriptive statistics about TSOVI and pause frequency across conditions are illustrated with the mean plot in Figure 3. The mean and confidence intervals are computed for the whole five-week period in the unit of group. In the DD condition, the indices are computed as the average values for all the participants in the group.

Figure 3(a) shows that, on average the DC groups spent more time on videos, and no noticeable difference is seen for the other two conditions. This finding is confirmed statistically with a mixed-effect linear regression analysis. A pairwise comparison revealed that the groups spent significantly more time in the DC condition than both in CC ($\beta=0.411$, 95% CI = $[0.143,0.679]$, $R^2=0.56$, $p<0.05$) and DD

($\beta=0.375, 95\%$ CI = [0.104,0.646], $R^2=0.56$, $p<0.05$) conditions. No significant differences were found between CC and DD conditions.

To complement the above finding, we compared the frequency of different types of video interactions (pauses and seeks). In terms of video seeks, the three conditions did not differ. However, the DC groups were found to pause significantly more often than in CC ($\beta=0.227$, 95% CI=[0.148,0.306], $R^2=0.49$, $p<0.005$) and DD ($\beta=0.214, 95\%$ CI= [0.133,0.295], $R^2=0.49$, $p<0.005$) conditions (see Figure 3(b)). Social effects can offer an explanation for this finding. The fact that each student in the DC condition was equipped with a mouse is perhaps seen by the students as a signal to encourage pausing videos. In CC condition, the single video-control made students hesitant to pause due to social pressures, such as when and who should make the pause. On the contrary, students in the DD condition, perhaps pause generally less in order to stay synchronized with each other, as we will discuss in the next section.

Synchronicity

In the scope of the DD condition, we are interested in how far individual students in a group synchronously watched videos together, which we denote as *synchronicity*. The synchronicity between two students in a group is obtained by computing the ratio between the total synchronous time and the length of the study session. Synchronous time means that the two students are either simultaneously watching the same video content or not watching any videos (i.e. they might be discussing). A threshold value *T* (measured in seconds) was introduced to determine the synchronous status. For each second of a study-session we look *T* seconds ahead and behind to see if the two students were or would be watching the same video content within *T*. In other words, we check if one student catches up with the other in "*T*" seconds. If yes, then they are synchronized. We coined the term **individual synchronicity index (ISI)** to measure the average synchronicity between all pairs involving the same student. Each student in a group has a different individual synchronicity, which signifies how the student synchronized with other students in the same group. Another term **group synchronicity index (GSI)** is used to denote the average of all individual synchronicities within a group.

An empirical value *T* = 50s was chosen to compute the previously defined synchronicity indices, because this threshold value maximizes the variance of resulting group synchronicity. The threshold also makes sense in real world: a teacher usually explains the same concept within this period, so it is reasonable to say that students are synchronized on the same ground.

Synchronicity over Time

The *GSI* over the five weeks for each DD groups (including the DSP group) are shown in Figure 4. The fifth week's data for the NAS-1 group is missing due to a technical problem during the experiment. This figure shows that the synchronization in groups is stable over time. In addition, a clear gap among the synchronicity series is seen in the middle range of the vertical axis, which separates highly synchronized groups (NAS-1 and NAS-2 groups with *GSI*>0.65) from lowly synchronized ones (NAS-3 and DSP groups with *GSI*<0.6).
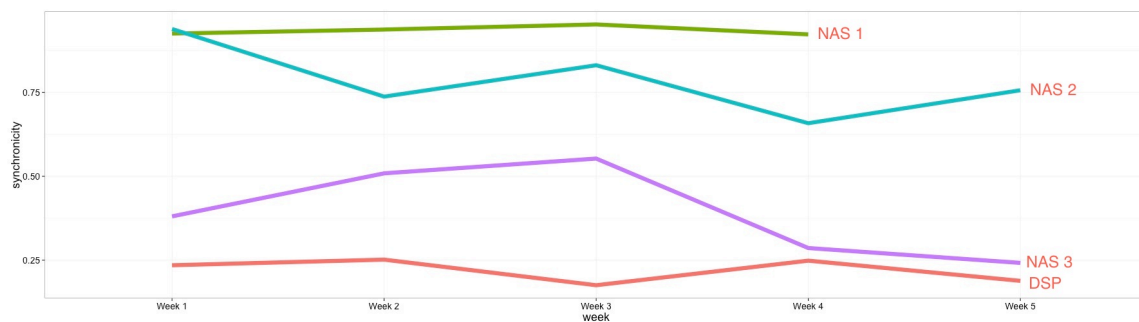
Figure 4. Changes in group-synchronicity index over weeks.

Variation in Synchronicity

While group synchronicities tend to stabilize over weeks, they vary across different groups. With mixed linear regression analysis, we found that the *TSOVI* is shown to have negative correlational effect with the *ISI* ($\beta=-0.19$, 95% CI = [-0.301, -0.078], $R^2=0.90$, $p<0.005$). This correlation indicates that more video-engagement time creates fewer opportunities for students in DD groups to keep synchronization. More time on videos suggests that many interactions (pausing and replaying) might have occurred, which apparently makes it difficult for students to stay synchronized. In the semi-structured interviews, highly synchronous groups reported that they usually noted down the problems while watching video-lectures, and discussed the problems right after everyone finished watching. The groups were self-regulated, and many students deliberately started and finished video watching more or less simultaneously.

We also asked the participants to rate their perceived equality of contribution among group members and the quality of their discussions on a 5-point Likert scale. We observed that highly synchronous groups perceived a high quality of discussion ($\beta=1.308$, 95% CI = [0.263,2.546], $R^2=0.445$, $p<0.05$), as well as equal contribution ($\beta=1.438$, 95% CI = [0.386,2.499], $R^2=0.35$, $p<0.05$). This result exhibits that synchronization is a desired attribute of a study group.
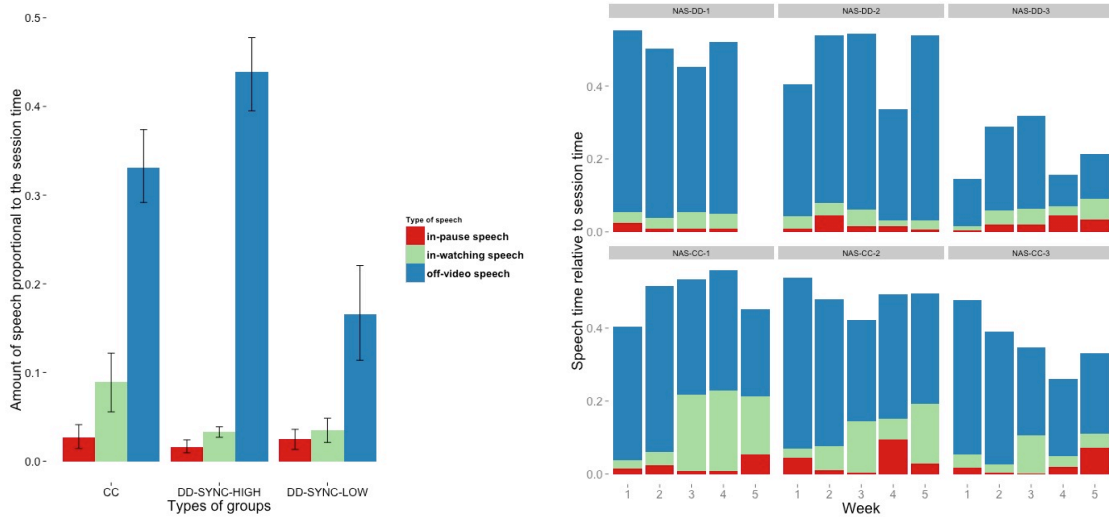
Amount of Speech

In the previous section we identified study groups in the DD condition with diverse synchronicities. DD groups are separated into two categories according to their synchronicity. In this section we investigate the effect of synchronicity on the amount of speech in the DD condition, with CC condition as the baseline for comparisons.

Speech Patterns

We categorize the speech during a study session into 3 types, ***in-pause speech***, ***in-watching speech***, and ***off-video speech***. The first type refers to the speech during the paused periods of time, the second concerns the speech while the videos are being played, and the last type represents the speech in the gap between watching two consecutive videos. The amount of each type of speech adds up to the total amount of speech in a study session. We measure the amount of speech at the group level, without differentiating whom it is from. The speech time is then divided by the total length of the corresponding session for normalization. The means of different types of normalized

speech are plotted in Figure 5(a) with confidence intervals, and detailed patterns for each group over the five weeks are presented in Figure 5(b). These graphs plot the speech data from the NAS DD groups only. The separation between highly and lowly synchronized group is consistent to that described in the previous section (NAS-1 and NAS-2: DD-SYNC-HIGH, NAS-3: DD-SYNC-LOW).

We first compare the common characteristics of different types of speech in Figure 5(a). The *off-video* speech contributed the largest to the total amount of speech. This type of speech happened after a video was finished, when students jointly solved quizzes and problems encountered in the video, both of which required group discussions. The fact that the amount of *in-watching* speech was always larger than in-pause speech is partly due to the overall higher length of video-play time compared to video-pause time. In addition, CC students could deliver spontaneous speech without pausing the video, and students were likely to discuss in subgroups, when some students were still playing videos. Both factors may also contribute to a higher amount of *in-watching* speech.

(a) Amount of speech among different conditions   (b) Amount of speech in each group over five weeks

Figure 5. Amount of different type of speech in study groups

Next we investigate the differences in speech across conditions (see Figure 5). The CC groups and highly synchronized DD groups resemble each other in large amount of total speech, whereas lowly synchronized DD groups overall talked noticeably less. It suggests that highly synchronized DD condition was almost like the CC condition in term of total speech, and the difference was that the CC groups talked more during video watching and less after the videos while the highly synchronized DD groups resulted oppositely. This observation can be explained by the group setup: The DD students were wearing headsets during video watching, so that they could not talk easily while watching videos. The loss of discussion during video watching was instead compensated after the videos. Note that even the speech patterns within the same group vary over five weeks (see Figure 5(b)). As for the comparison within DD condition, the interpretation of the bar charts for the lowly synchronized groups needs to be made with caution, since only one group is classified in their category. To get a better picture of the dynamics in the DD condition, we use statistical tools to quantify the patterns, which will come next.

Effect of Synchronicity on Speech

Figure 5(a) shows distinct patterns for lowly synchronized DD groups, which suggests a potentially systematic effect of synchronicities. With mixed linear regression analysis, we found that synchronicity (ISI) had a positive correlational effect on the amount of *off-video* speech ($\beta=0.45$, 95% CI = [0.319,0.621], $R^2=0.84$, $p<0.0005$) and a negative effect on the amount of *in-pause* speech ($\beta=-0.06$, 95% CI = [-0.076, -0.011], $R^2=0.21$, $p<0.05$). This result suggests that more synchronized groups spent less time in pauses within videos, but more time after watching the videos. To complement the result above by relating the speech to video interactivity, we found that video interactivity (TSOVI) positively affects *in-pause* speech ($\beta=0.05$, 95% CI = [0.026,0.074], $R^2=0.28$, $p<0.0005$), and negatively affects *off-video* speech ($\beta=-0.113$, 95% CI = [-0.202, -0.022], $R^2=0.82$, $p<0.05$). This result is in line with the negative correlation between ISI and TSOVI. The interactions between these correlations are shown in Figure 6.
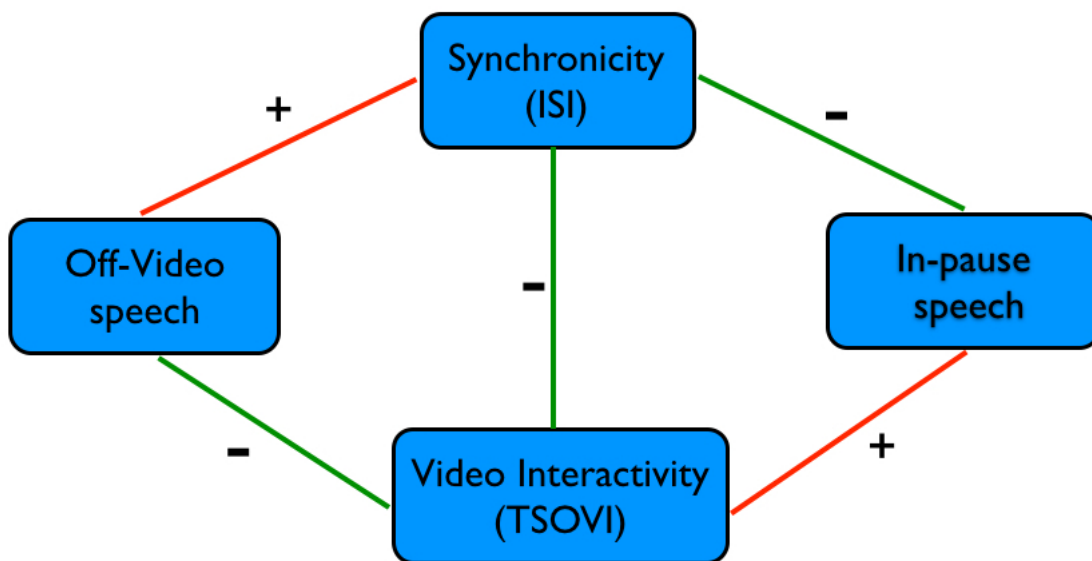

Figure 6. Relationship between video interactivity, synchronicity and amount of speech

Figure 6 suggests that the more students in the DD groups engaged in video watching, the less synchronized they were, which in turn increased the amount of off-video speech and decreased the in-pause speech. In other words, highly synchronized students sacrificed video engagement for gaining synchronicity, resulting in more discussions during off-video periods. The amount of off-video speech seems to be balanced with that of in-pause speech: one aspect wanes, the other waxes. This effect is confirmed with marginal significance ($p=0.07$). However, the link between them is missing, because the amount of off-video speech is affected by the time spent on quizzes as well. It is difficult to conclude there is a relation between these two types of speech

Discussion

Although previous MOOC research (Mackness, Mak & Williams, 2010) suggests that learners have tendency towards group learning, our study puts it into practice within a collocated setting. Our experiment participants are authentic MOOC students, the

conditions are quasi-realistic and the study has the ingredients to bring the study-group way of learning with MOOCs in the mainstream.

First and foremost, our observations show that the study-group way of watching MOOC videos was widely accepted and highly appreciated by the learners. The study participants reported on increased levels of attentiveness and engagement. In addition, learning within study-groups can render the learning experience with difficult courses a pleasing experience, and helps the group-members to validate each other's understanding of the topic; resulting into a shared knowledge model. Furthermore, a high quality of discussions was also perceived by the participants regardless of the experimental condition, which conforms to the previous statement that being in a group is crucial towards learning with MOOCs. These observations indicate that our study can have implications towards flipped-courses within the universities as well as towards promoting learning with MOOCs in infrastructure-scarce developing countries.

Next, we observed that watching videos synchronously (denoted by synchronicity) emerged as a crucial attribute of the study-groups. The two experimental conditions with the centralized display (CC and DC) were already synchronized as all the group members saw the same part of the video. However, we found interesting results in the DD condition where the group members watched the videos on their own. We observed a split in this condition, as some groups chose to stay highly synchronized by sacrificing the power of interactivity with the video. These highly synchronized groups discussed significantly more between videos as compared to the lowly synchronized groups, and were at par with one of the synchronized condition with centralized display. In addition, highly synchronized groups perceived discussions to be of better quality. These observations indicate that synchronicity while watching videos empowers learners with a sense of being in a *team*. On the contrary, interacting too much with the video promotes individualism, as the learners are only concerned with their own learning rather than sharing and validating their understanding with others; and this idea also goes against the theme of study-groups.

Further, co-located video watching also empowers study-group members to discuss and validate each other's understanding of the topic, whenever faced with confusion or a difficulty. This situation is analogous to asking questions to a teacher in the middle of a lecture, with the only difference that in a study-group one can ask questions to her peers. Therefore, presence of individual controls is advantageous as it promotes more discussions while watching videos. This is made evident as the groups in DC condition were observed to be more interactive (with the video) and spent more time on videos as compared to CC and DD conditions. One possible explanation for this difference could be that in CC condition the group members were hesitant to reach out and explicitly acquire the video control, and possibly only dominant members did so. On the other hand, the lack of awareness of what the others were watching in the highly synchronized DD groups led to less interactions with the videos. This observation is in line with the benefits of single display groupware in peer-tutoring and peer-learning as identified by Stewart et al. (1999). Single display groupware, such as the one used in DC condition, increase awareness of the learning activity and group members, promote balanced participation, and mitigate the chances where one group member takes explicit control of shared resources (video control). This leads us to the conclusion that DC condition is a good fit for situations where high interactivity with the video lectures is desired while keeping the synchronicity intact.

In this study, as our focus was on understanding the group-study patterns that naturally emerge under different conditions, we did not measure the learning outcomes of the study-group participants. However, this could be an interesting research question

for future work, as it will enable us to identify the optimal condition for learning with MOOCs in collocated settings. Moreover, the lack of comparison between group and individual video watching adds a limitation to our study, as we cannot conclude the relative value of group versus individual learning with MOOCs.

Conclusion

Co-located MOOC watching has the potential to become widespread at universities as well as geographically centered clusters, both due to the large learners base and the simplicity of technical solutions. A simple study group configuration requires no more than conventional computers, as demonstrated in this article. We have investigated how spontaneous MOOC study groups watch videos together under different configurations. Our results reveal that watching MOOCs in groups provides highly satisfying learning experience as learners feel connected and interactions among them are enabled. We identified that synchronicity while watching video lectures is a desirable attribute of collaborative video watching. However, when both video control and display were distributed, students tended to sacrifice in-video discussions and video interactions for staying synchronized. A video watching condition with shared display and distributed individual controls might enable study groups to find a fine balance between synchronicity, video interactivity and discussions. These results have implications for MOOC pedagogical designers towards the integration of collaborative learning practices in the existing MOOC platforms.

References

Brindley, J., Blaschke, L. M., & Walti, C. (2009). Creating effective collaborative learning groups in an online environment. *The International Review of Research in Open and Distance Learning*, *10*(3).

Cadiz, J. J., Balachandran, A., Sanocki, E., Gupta, A., Grudin, J., & Jancke, G. (2000, December). Distance learning through distributed collaborative video viewing. In *CSCW'2000* (pp. 135-144). ACM.

Capdeferro, N., & Romero, M. (2012). Are online learners frustrated with collaborative learning experiences?. *The International review of research in open and distance learning*, *13*(2), 26-44.

Gibbons, J. F., Kincheloe, W. R., and Down, K. S. (1977). Tutored videotape instruction: a new use of electronics media in education. *Science*. 195: 1139-1146.

Johnson, D. W., & Johnson, R. T. (1989). *Cooperation and competition: Theory and research*. Interaction Book Company.

Kop, R., Fournier, H., & Mak, J. S. F. (2011). A pedagogy of abundance or a pedagogy to support human beings? Participant support on massive open online courses. *International Review of Research in Open and Distance Learning*, *12*(7), 74-93.

Mackness, J., Mak, S., & Williams, R. (2010). The ideals and reality of participating in a MOOC. In *Networked Learning Conference* (pp. 266-275). University of Lancaster.

Martin, F. G. (2012). Will massive open online courses change how we teach?. *Communications of the ACM*, *55*(8), 26-28.

Sipusic, M., Pannoni, R., Smith, R., Dutra, J., Gibbons., J., and Sutherland, W. (1999). Virtual Collaborative Learning: A Comparison between Face-to-Face Tutored Video Instruction and Distributed Tutored Video Instruction (DTVI). Sun Microsystems Laboratories, Inc. TR-99-72. Available at http://www.sun.com/research/techrep/1999/ abstract-72.html

Smith, R. B., Sipusic, M. J., & Pannoni, R. L. (1999). Experiments comparing face-to-face with virtual collaborative learning. In *CSCL'99* (p. 68). International Society of the Learning Sciences.

Smith, G. G., Sorensen, C., Gump, A., Heindel, A. J., Caris, M., & Martinez, C. D. (2011). Overcoming student resistance to group work: Online versus face-to-face. *The Internet and Higher Education*, *14*(2), 121-128.

Siemens, G. (2005). Connectivism: A learning theory for the digital age.*International journal of instructional technology and distance learning*, *2*(1), 3-10.

Siemens, G. (2006). *Knowing knowledge*. Lulu.com.

Stewart, J., Bederson, B. B., & Druin, A. (1999). Single display groupware: a model for co-present collaboration. In *CHI'99* (pp. 286-293)*.* ACM.

Stone, H. R. (1990). Economic development and technology transfer: Implications for video-based distance education. *Contemporary issues in American distance education*, 231-242.

Tang, K. C. C. (1993). Spontaneous collaborative learning: A new dimension in student learning experience?. *Higher Education Research and Development*, *12*(2), 115-130.

Tucker, B. (2012). The flipped classroom. *Education Next*, *12*(1), 82-83.

Vygotskiĭ, L. L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard university press.

Weisz, J. D., Kiesler, S., Zhang, H., Ren, Y., Kraut, R. E., & Konstan, J. A. (2007). Watching together: integrating text chat with video. In *CHI'2007* (pp. 877-886). ACM.