Seeing the Instructor in Two Video Styles: Preferences and Patterns

Suma Bhat
University of Illinois
Urbana-Champaign, USA
spbhat2@illinois.edu

Phakpoom Chinprutthiwong
University of Illinois
Urbana-Champaign, USA
chinpru2@illinois.edu

Michelle Perry
University of Illinois
Urbana-Champaign, USA
mperry@illinois.edu

ABSTRACT

Instructional content designers of online learning platforms are concerned about optimal video design guidelines that ensure course effectiveness, while keeping video production time and costs at reasonable levels. In order to address the concern, we use clickstream data from one Coursera course to analyze the engagement, motivational and navigational patterns of learners upon being presented with lecture videos incorporating the instructor video in two styles - first, where the instructor seamlessly interacts with the content and second, where the instructor appears in a window in a portion of the presentation window.

Our main empirical finding is that the video style where the instructor seamlessly interacts with the content is by far the most preferred choice of the learners in general and certificate-earners and auditors in particular. Moreover, learners who chose this video style, on average, watched a larger proportion of the lectures, engaged with the lectures for a longer duration and preferred to view the lectures in streamed mode (as opposed to downloading them), when compared to their colleagues who chose the other video style. We posit that the important difference between the two video modes was the integrated view of a 'real' instructor in close proximity to the content, that increased learner motivation, which in turn affected the watching times and the proportion of lectures watched. The results lend further credibility to the previously suggested hypothesis that positive affect arising out of improved social cues of the instructor influences learner motivation leading to their increased engagement with the course and its broader applicability to learning at scale scenarios.

1. INTRODUCTION

Lecture videos constitute the primary source of course content in the massively open online courses (MOOCs) offered by platforms such as Coursera and EdX. Not surprisingly they are also the most-used course component (compared to quiz submissions and discussion forum participation)[4, 12, 17]. Owing to the asynchronous and virtual nature of teaching and learning in these environments, lecture videos comprise the only channel through which learners have access to their instructors, an important factor affecting student motivation, satisfaction, and learning [19].

The important role of lecture videos as the primary content-bearers of a course results in instructional content designers rightly concerned about optimal video design guidelines that ensure course effectiveness; of having video lectures that maximize student learning outcomes while keeping video production time and costs at reasonable levels [9].

A recent study addresses some aspects of these concerns by comparing learner engagement patterns with video lectures across courses in the context of MOOCs [9]. The outcome of the study was a set of broad recommendations answering the concerns at a broad level. In particular, one of the take-away messages was to include the instructor's head in the presentation at opportune times by means of a picture-in-picture view of the instructor. From the perspective of this past work, our current study is a more focused version of [9]. Using the case of a Coursera course that *concurrently* made its video lectures available in two modes (the modes differ in ways in which they present a view of the instructor), the current study is unique in that it seeks to refine the recommendations made in [9]. We do this by observing how learners interact with the course in a MOOC-sized community. The central component of the current study is an empirical analysis of the course logs to highlight the differences and similarities between the motivational, navigational and engagement tendencies of the users who interact with the two available lecture modes. The uniqueness of the study is that the same set of lectures is available in two modes, which permits us to see if there are navigational behaviors and engagement patterns that are supported by specific video types.

Our empirical findings in this study are summarized below: When comparing users who watched the lectures in only one video mode,

1. We observe that learner group preferences of one mode over the other differ considerably with a ratio of 10:1.

2. Learner group preferences of the video mode for viewing lectures directly translate to differences in the pro-

portion of available lectures watched, engagement times with the videos (via differences in watch times) and in the manner in which videos are watched (streamed vs. downloaded) between the two groups.

3. Certificate earners and auditors (learners who primarily engage with a course by only watching videos) were more likely to choose one video mode over the other.

In addition, analyzing users who watched video lectures in both modes (switching twice - from one mode to the other and back to the mode first used),we notice that the disparity in preference persists (as noted above in the case of users who watched only one video mode), although the within-user differences in engagement times and the proportion of lectures watched were not statistically significant.

While many factors could be at play here, and while proposing the need for further studies to confirm our hypothesis, we posit that the video mode preferred by the majority of learners who use only one mode has the following advantage; it offers an integrated, rather than separated, access to the instructor's eye-gaze (whether the instructor is looking at the student or the content) and gestures in close proximity to the lecture content that results in a better learning experience for the learners via the availability of more realistic social cues.

2. RELATED WORK

MOOCs are criticized for their high attrition rates and are alluded to as a learning environment where a majority of students are passive lurkers who do not actively engage with the course. The low levels of engagement and completion could, in part, be attributed to the demand of the MOOC environment. MOOCs require students to be autonomous learners, who can remain motivated despite low levels of instructor presence in the course, the feeling of isolation and the unclear sense of purpose in an asynchronous learning environment. Unfortunately, aside from a handful of interactions in online discussion forums, the pre-recorded videos are the only chances for an instructor to create a sense of presence in a MOOC environment.

Prior analyses of MOOCs (e.g. [4]) have found that students spent the majority of their time watching lecture videos and that many students are auditors whose course interaction is limited primarily to watching video lectures [12]. It then follows that the design of effective videos is a critical component not only for learning effectiveness but also for the success of the course in terms of making the material accessible not just to certificate earners but also to auditors.

The design of effective video lectures, however, is informed by studies in psychology, cognitive science and online learning. Recent findings suggest that a richer instructor-student interaction in an online course is afforded by video-based sessions when compared to courses with only audio narration [3]. In addition, studies on online learning reveal that learners need to have a sense of relatedness to their instructors and that this sense is often communicated through information that is superfluous to the learning objectives [19, 5]. For instance, the presence of a humanoid pedagogical agent, be it in the form of an avatar or a cartoon figure, in a computer aided learning environment can improve a student's learning experience [6].

While the importance of non-verbal modalities of interaction (via gestures and eye-gaze) in human-human communication has long been recognized [18, 1], only recently are non-verbal modalities being harnessed in virtual communication scenarios (e.g., access to the course instructor in a window at the corner of the presentation screen in a video lecture). It is likely that increasing access to non-verbal communication can improve the instructor's sense of presence in an online-only learning environment such as a MOOC, and thus improve students' learning and their desire to stay engaged in their learning.

Clark and Mayer [6] emphasize the effectiveness of bringing instructor non-verbal modalities to the presentation because they encourage deeper engagement with the lecture content and trigger social responses in the learner [16, 7]. However, empirical evidence on its effect on learning outcomes is largely inconclusive [14, 15].

The effect of the instructor's face in visual attention, information retention and learner affect has been explored in studies such as [11, 2]. In [11] it was found that including an instructor's face in a presentation resulted in positive affective response in learners which in turn influenced the time devoted to learning. However, access to the instructor's face had no specific effect on attention or retention. In [2], an analysis of the perceptions of students being presented with two modes of video lectures incorporating the instructor's face in the presentation is available. Results suggested that having access to the instructor's gestures were potentially related to increased user satisfaction. Both these studies were not conducted in MOOC-scale environments and had a small subject pool ([11] had n=22, and [2] had n=60).

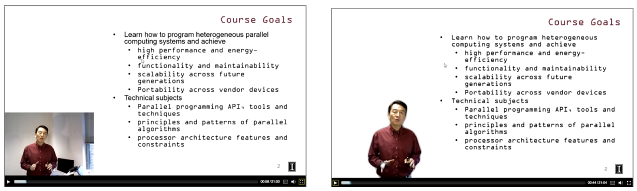
In [9] the results of a retrospective study based on course logs of MOOCs showed the effect of different video lectures produced in different styles on the engagement patterns of learners. Based on a large dataset, results indicated that video lectures that involved a talking head were more engaging to the students than lectures without a talking head. The recommendation based on these results was to include the instructor's head in the presentation at opportune times by means of a picture-in-picture view of the instructor.

This study is set with a similar goal such as that of [9] - that of understanding learners' navigational and engagement patterns with different modes of video presentations. The different modes are chosen in a way that afford access to the instructor as recommended in [9]. This permits us to see if there are navigational behaviors and engagement patterns that are supported by specific video types.

Three factors set this study apart from prior related studies. First, we compare two modes of lecture videos with access to the instructor in the *same* course. Second, the two video modes are available to the learners over a reasonable duration (three weeks/22 lectures) thus permitting the analysis over a longer duration compared to studies [11] and [2]. Third, the setting is a realistic learning at scale setting where students rely solely on video instruction.

3. METHOD

We conducted a *retrospective study* of the engagement, motivational and navigational patterns of learners as a response to video lectures presented in two styles. The learners were enrolled in the Coursera course on programming massively parallel processors offered from January to March 2014.



(a) PiP mode (b) Overlay mode

Figure 1: Screenshots of the two video modes of the lectures

3.1 Video Styles

Today's advancement in video capture technology allows for ways of improving an instructor's presence in the online classroom by including the instructor's face in the presentation at substantial reductions in video production costs. The video lectures for the course were available in two modes: the *picture-in-picture* mode and the *overlay* mode both produced in non-studio settings by the instructor and recorded simultaneously. The audio quality for both modes was excellent and similar.

Picture-in-picture mode: Presentation creation technologies can embed a video of the instructor inside a presentation, with the instructor appearing inside a window alongside the content window. In this course, the instructor window appears in the lower left corner of the presentation. We will refer to this video style as the *PiP* mode (see Figure 1a for a screenshot of this mode). The size of the instructor's video is limited by the constraints of window placement in the presentation screen.

Overlay mode: New screen capture tools are able to capture only the instructor's video without the background and overlay the video of the instructor into a presentation such as PowerPoint slides much like the green screen technology used in weather forecasts. As a result of this overlay and the screen capture technology, the instructor is able to interact with the content seamlessly by pointing at relevant sections via gestures. In addition, the instructor appears in a much closer proximity to the content window, and in a larger relative proportion compared to the instructor appearing in a window alongside the content window (PiP mode above).We will refer to this video style as the *overlay* mode (refer to Figure 1b for a screenshot). Notice how the instructor appears beside the content on the left.

The first 22 lectures, which constituted the material of the first three weeks of the course, were offered in these two modes. Both modes were available in the video lectures page on the course wiki during the entire duration of the course and were available for streamed view as well as for download. The average duration of the videos was 19.23 min. The file size of a lecture in overlay mode was about 1.2 times that of its corresponding PiP version. When the course began the course syllabus had a note about the availability of the lectures in two modes for the first three weeks and that the students were free to choose the format of their choice.

Because this was a retrospective study and not a controlled study, rather than assigning users to watch a given mode, we observed how students used the resources and interacted with them. The users[1] were classified into three groups based on the lecture modes they viewed (a user who clicked to view at least one lecture was counted in the group). There were users who viewed the lectures of the first 3 weeks only in the PiP mode (we call this group the **PiP** group, N = 899), those who viewed them only in the overlay mode (we call this group the **Overlay** group, N = 5740) and those who viewed them in both modes (the **Both** group, N = 3791). We compare the groups with respect to the analysis variables described below.

3.2 Analysis Variables

We created the following sets of analysis variables to reflect aspects of engagement, motivation and navigation.

Engagement: Because our analysis was based on the course logs, a true measurement of learner engagement is impossible. We approximate engagement via two proxy measures:

<u>Video watching time (**wtime**)</u>: This is the total length of time that a student spends viewing video lectures (lectures 1 to 22) and we use it as the main index of engagement. This measure is limited in scope because it only provides information for streamed lecture views. Moreover, it has no indication whether the engagement with the video is an active one or a passive one (as in playing it in the background).

<u>Discussion forum visits following a lecture view (**dfvisit**)</u>: We use a visit to the discussion forum (either to begin a thread, comment on an existing post or view a related post) immediately following a lecture (within 30 minutes) as an index of engagement. This reflects the intent of the learner to be open to aspects of the lecture beyond what is available in the video lecture.

Motivation: A limitation of this retrospective study was that access to learners' motivation (by interviewing a sample of learners, for instance) was unavailable. As a proxy to measuring motivation, we consider the following two indices:

<u>Certificate-earner proportion (**certprop**): The fraction of users who went on to earn a certificate.</u>

<u>Coverage (**cov**)</u>: The fraction of lectures (and quizzes) that the learner viewed (and submitted) is our second measure of motivation. Again, an important limitation of this measure is that it only represents the fraction of lectures viewed in the streamed mode and gives no indication about those viewed after downloading[2].

Navigation: We analyzed the navigation behavior of the

[1]We only took into account users who did not explicitly drop the course.

[2]Analysis of this variable by limiting it to users who only watched a video streaming would have been a possibility but for the fact that the sample for PiP was very small (< 30).

students by observing their interaction with the course components. The measures we use are:

Streaming index (**SI**): In [12] streaming index was used as a measure of video consumption and is defined as the proportion of overall lecture consumption that occurs online on the platform (streamed), as opposed to off-line (downloaded),

$$\text{Streaming Index(SI)} = \frac{\text{streamed lecture consumption}}{\text{total lecture consumption}}.$$

Here we use it as a measure of video access.

Discussion forum activity (**dfview** and **dfpost**): The discussion forum constitutes a highly under-utilized resource in a MOOC platform and activities associated with it can be considered to be an important index of interaction with the course. Even though this measure involves a minority of course participants, we compared the number of views and posts by the users in the two groups to see if users of a video group show a tendency to participate more in discussion forums.

Back-jump proportion (**bjprop**): As used in [10], we first define a learning sequence as an ordered sequence of learning activities and its length as the number of activities in the sequence. An example of a learning sequence of length two in one session would be a lecture view followed by a quiz attempt. For our study, we consider the learning sequences of the users involving the first 22 lectures and the associated quizzes limiting the learning activities to lecture views, quiz attempts and quiz submissions.

A back-jump is a backward navigation in a learning sequence. The count of back-jumps indicates the number of times a student navigated backwards in the learning sequence and is suggestive of a departure from a linear learning sequence. In our case, this would be from a lecture to a lecture release earlier (lecture 4 to lecture 2) or from a quiz to a previous lecture (such as quiz 3 to lecture 2.3). Back-jump proportion is the number of back-jumps divided by the length of the learning sequence of the student. In [10], this measure served as an index of non-linear navigation through the course material to differentiate field-dependent learners (those who follow a sequential learning path as laid out by the content creators) from field-independent learners (those who resort to a non-linear fashion of exploring the learning environment) [8, 13], which we use in our study as well.

Other measures of comparison such as that of performance (in terms of quiz scores and assignment scores) could have been used here, but the course managed them in a server whose logs were not available in the Coursera data set.

4. EMPIRICAL OBSERVATIONS

The groups **PiP** and **Overlay** (as described in Section 3.1) are first compared with respect to the analysis variables just described and the resulting observations are summarized. Following that we analyze the users in the **Both** group.

We chose a course-week (as listed in the course wiki) as a unit and counted the number of video views during that week. In Figure 2 we see the number of unique views by the users in each of the groups during the first 3 weeks. Each

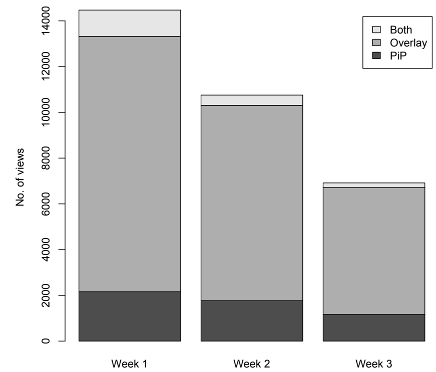


Figure 2: The number of video views in each group (Overlay, PiP and Both) over the first three weeks of the course.

bar includes the number of unique views of all lectures by a particular group during that week. What is apparent from the figure is that, over the three weeks when the lectures were available in two modes, a majority of views occurred in the Overlay mode. In addition, it is of interest to note that even in the third week there was a non-trivial number of users who watch both the modes. These views could be attributed both to the late entrants to the course and to those who switched modes in that week.

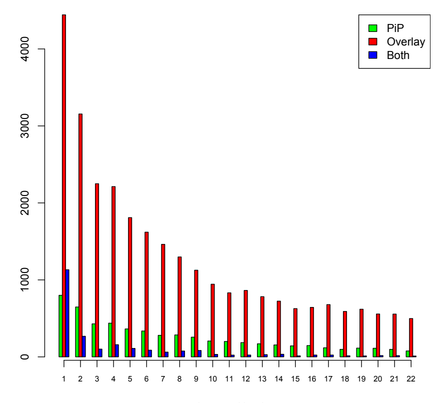


Figure 3: No. of views of each lecture over the duration of the course.

Another perspective of the views of each group is available in Figure 3 which shows the number of unique views of the 22 lectures by users in each group. Here again we notice that the *Overlay* mode was preferred by the vast majority of users compared to the *PiP* mode. It is also interesting to note from Figure 3 that the number of users who viewed the lectures in both modes is quite significant (even larger than the number of views in the *PiP* mode) for lecture 1 and then drops drastically for the lectures that follow. This could be interpreted to mean that users decide on their preferred mode as early as the first lecture. (In both these plots, the decrease in the number views is indicative of learner attrition through the duration of the course.)

4.1 Analysis Variables Compared

We filtered out all users whose total watching time lasted less than 110s (approximating individual sessions lasting on an average shorter than 5s which could have been a result of users who paused immediately after beginning to watch a video or navigated to another page). This resulted in groups of size 385 (**PiP**), 3725 (**Overlay**) and 3791 (**Both**) respectively. Below we summarize the results upon comparing the analysis variables between the fist two groups.

A majority of the analysis variables considered here have highly skewed distributions thus deviating from the assumptions of normality. Under these circumstances, we resort to the Mann-Whitney U test to compare the two distributions. The null hypothesis tested here is not that the medians (or means) are equal but that the two groups come from the same underlying distribution. That is to say, we are testing for equality of location and shape of the distributions, not for equality of any one aspect of the distribution. Although the distributions were skewed we tabulate the mean of the variable for the two groups for the purpose of representation (see Table 1. The final column of the table indicates the p-value of the Mann-Whitney test. Statistically significant differences between groups are indicated in bold-face.

The *Overlay* and the *PiP* group: From Table 1, we observe that the underlying distributions for watch time, coverage, and streaming index differs significantly between the two groups. The **Overlay** group had a larger mean watch time compared to the **PiP** group (median watch times=33.65 min. and 21.55 min. respectively). In addition, streaming is the dominant way of accessing videos for both the groups. Streamed videos constituted an average 77% of the video usage for the **Overlay** group as opposed to 60% for the **PiP** group (respective medians 93% and 66%).

Measure	Overlay	PiP	p-value
Watch time (min)	**83.82**	**63.32**	< 0.01
Disc. forum visit	0.29	0.24	0.23
Certificate prop. (%)	8.48	6.75	0.24
Coverage	**0.24**	**0.18**	< 0.01
SI	**0.77**	**0.60**	< 0.01
Forum post	0.36	0.43	0.80
Forum view	11.86	17.22	0.59
Back-jump prop.	0.09	0.09	0.92

Table 1: Comparison of the measures for the two groups.

The 95% confidence interval of the two medians for wtime were (26.64, 38.75) for *PiP* and (49.77, 55.46) for *Overlay*. For SI the 95% confidence interval of the two medians were (0.8332, 0.8333) for *Overlay* and (0.564, 0.649) for *PiP*. Because the two confidence intervals for the medians of each group were non-overlapping, we infer that the corresponding distributions are different (also indicated by the Mann-Whitney U test).

This situation lends itself to two possible interpretations. Either more videos were watched streaming (with the same number of downloaded videos), or more Overlay videos were streamed compared to PiP with fewer Overlay videos downloaded. Both the interpretations imply that the streamed view was the primary way in which videos in Overlay mode were accessed.

As for coverage, we found that users in the *Overlay* group viewed a larger proportion of available lectures compared to their colleagues in the *PiP* group. Taken together with the lower coverage for *PiP*, its lower watch time is then justified since a smaller proportion of video views were streamed.

Although we noticed an apparent difference in the proportion of certificate earners between the two groups, a two-sample Z-test indicates that the difference in proportion was not statistically significant (p= 0.24).

Certificate Earners: We next restricted the analyses to the certificate-earners of the course, knowing that these were the most committed users in a course. The results limited to the certificate earners (N=316 for Overlay and 26 for PiP) are summarized in Table 2.

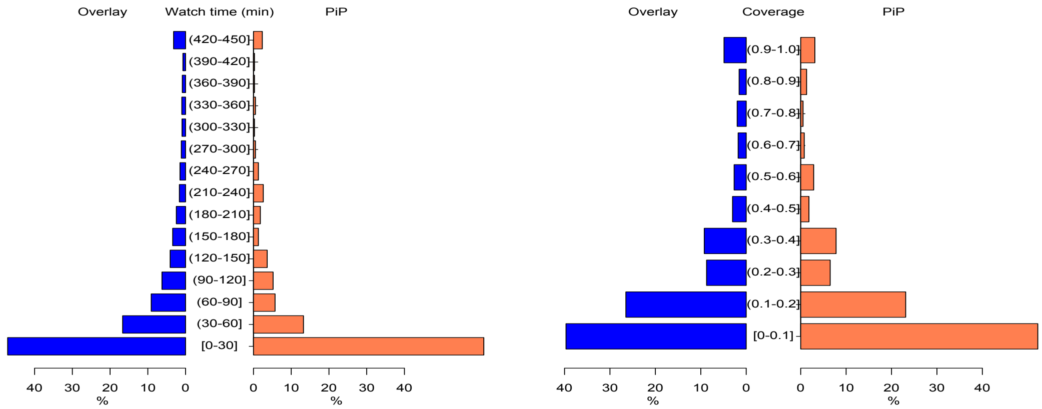
Measure	Overlay	PiP	p-value
watch time (min)	233.35	194.57	0.23
Disc. forum visit	1.53	1.69	0.84
Coverage	**0.70**	**0.58**	< 0.01
Streaming Index	**0.70**	**0.56**	0.02
Forum post	2.25	3.23	0.18
Forum view	76.44	113.08	0.08
Back-jump prop.	0.09	0.05	0.12

Table 2: Comparison of the measures for certificate earners.

We first computed the posterior probability of a certificate earner choosing one video mode over the other. Using empirical counts, we have the priors of the three groups: the probability of choosing the Overlay mode is 47%, that of choosing *PiP* is 5% and that of choosing *Both* is 48%. We also have the likelihoods: the probability that the student is a certificate-earner given that the student chose Overlay is 8.5%, the probability that the student is a certificate earner given that the student chose PiP is 6.8% (both from Table 1) and the probability that the student is a certificate-earner given that the student chose Both is 10.4% (empirically obtained).

Using this information, we calculated the probability that a certificate-earner chooses Overlay to be 0.43, that he/she chooses PiP is 0.04 and that he/she chooses Both is 0.53. This suggests that that a certificate earner is most likely to try both before settling for one mode. However, among the two modes, the more likely choice would be the Overlay mode.

Limiting the comparative analysis to the certificate earners of the two groups, from Table 2 we notice that the trends observed in the overall comparison are also largely applicable here with the exception of watch time. A surprising observation here is that despite the differences in the distributions for coverage and streaming index, differences in the distributions of the video watching times were not statistically significant. A likely explanation is that the certificate earners in the PiP group revisited portions of the same video, resulting in longer watch times compared to their Overlay



(a) Distribution of Watch times

(b) Distribution of Coverage

Figure 4: Histograms of Watch time (left) and Coverage (right) for the two groups compared. Each plot shows the density corresponding to each bin in the y-axis.

colleagues.

What is new here is that certificate earners in the *Overlay* group show apparently different non-linear navigational patterns compared to their *PiP* counterparts as evidenced by the difference in means. However, the distribution of back-jump fractions is not statistically significant (p=0.12) possibly owing to the relatively small sample size of the PiP certificate earners (n=26).

Auditors: From [12] we know that auditors (defined in that study as learners who did assessments infrequently if at all and engaged instead by watching video lectures) are nearly as engaged and motivated in the course as certificate earners in terms of using lecture materials in MOOCs and show similarly high levels of overall learning experience to certificate earners. Here, we investigate the extent to which users in the two groups had engagement levels similar to that of certificate earners.

We identified the auditors by clustering the users using k-means in the Overlay and the PiP groups by three factors into 3 classes (certificate earners, auditors, and lurkers):

- coverage (answering the question 'How many lecture units were watched?');
- streaming index (answering the question 'How were the lectures watched?');
- watch time (answering the question 'For how long were the lectures watched?').

We observed that the certificate users fell into a predominant group, which also included a set of non-certificate users 'similar' to the certificate users; these users behaved like the certificate users with respect to the 3 factors considered here. We refer to these users as auditors since they used resources much like the certificate users, except for the fact that they did not earn a certificate. We noticed that 3.5% of Overlay users were auditors in this sense and nearly 6% of users in the PiP were auditors. The difference in proportion of auditors was statistically significant (p=0.012), suggesting that PiP had a larger proportion of auditors compared to Overlay.

We then calculated the likelihood of an auditor choosing a specific viewing mode using empirical counts and note that the probability that an auditor chooses Overlay was 0.86 much greater than the probability that an auditor chose PiP, which was 0.14.

4.2 The Both group

While a comparison between the Overlay and the PiP groups served as a type of between-subjects analysis, a within-subjects type of analysis is afforded by analyzing the Both group. Although users watched both video modes in this group, to get a more reliable picture of engagement patterns and video mode choices, we included only those users who watched at least half of all the available lectures. With this set-up we assume that the users had sufficient exposure to the mode in which they began watching lectures before switching to the other mode. In addition, they had sufficient opportunities to experience the second mode and revert back to the original mode if they chose to do so.

Users in this group watched lectures in both modes and could be divided into three groups: 1) those who viewed a set of lectures in one mode and then switched to the other mode and remained in that second mode for the rest of the lectures, 2) those who switched twice eventually returning to watch the remaining lectures in the original mode in which they began, and 3) those who showed no apparent preference for one mode over another. For the purpose of our analysis, we focus on the second of these three groups because the sample size of the first group was too small (< 30) to draw meaningful inferences and we had no meaningful analyses to conduct with the third group.

With this restriction on the users, we were left with 271 users (34% of the users in Both), of which 241 (89%) watched

	OPO	POP	p-value
Coverage	0.71	0.61	< 0.01
Streaming Index	0.80	0.57	< 0.01
Watch time (min)	291.69	260.85	0.10
Disc. forum visit	1.83	1.61	0.56
Back-jump prop.	5.6	4.6	0.15
Certificate prop.	0.37	0.63	<0.01

Table 3: Comparison of the mean values of the measures for the users in the *Both* group.

most of the lectures in the overlay mode and the remaining 30 watch most of the lectures in the PiP mode. It is clear that the majority of users in this group began watching the lectures in the overlay mode, switched to the PiP mode, and reverted to watching in the overlay mode. We represent this majority group as OPO and the other group as POP. For each user in the POP and OPO groups, we computed the measures of coverage, streaming index and watching time over the lectures watched in a given mode, yielding a measure for each video mode watched. We summarize these measures in Table 3.

We observe from Table 3 that the distributions of coverage and streaming index for the Overlay mode and PiP mode differ substantially and that the difference is statistically significant. We infer that a larger proportion of lectures were watched by the users following an OPO pattern compared to a POP pattern and that the videos in Overlay mode were streamed, while the videos in PiP mode were mostly downloaded. We we notice that the distributions of watch times were not different between the OPO and POP. This implies that when the users had a chance to watch both the modes, their engagement patterns with their 'preferred' mode was similar.

Unlike in the case of the groups that watched only one mode, a comparison of the proportion of certificate earners between the two Both groups shows that a larger proportion of POP were certificate earners and that the difference in proportion was statistically significant via a two-sample Z-test ($p < 0.01$).

5. INTERPRETATION OF RESULTS

The present study suggests that learners showed a strong preference for the Overlay mode over the PiP mode. Comparing the user groups that viewed the lectures in only one mode, we saw that the two groups differed significantly in their watching times, choice of video access and proportion of lecture materials viewed. The preference of Overlay was also exhibited by the users that watched both modes. This suggests that the Overlay mode was preferred and we hypothesize that these videos appeared more engaging. Taken in light of the results of studies such as [7], the findings here could be interpreted to mean that this was the result of a positive affective response of the learners to social cues in the learning environment (here the videos). It is likely that the overlay mode offered several affordances over the PiP mode – integrated rather than separated access to the instructor's eye-gaze and gestures, the instructor's proximity to the slides, and the larger size of the instructor – which could have yielded differences in social cues available via the video modes.

This primary social cue that was different between the two video modes, we hypothesize, was the integrated view of a real instructor and this is likely to have increased learner motivation, which then affected the amount of time learners spent watching a lecture and the proportion of lectures they watched. Aside from this hypothesis on the difference in the availability of social cues, in the absence of watching actual behaviors of the learners affording a more fine-grained characterization of their watching patterns (such as the actual time users spent watching the video or the amount of time they spent looking at the instructor's face) and a qualitative analysis via interviewing users for their opinions about the videos, the true implications of the difference on the video watching/consuming patterns cannot be determined. Another set of experiments to quantify the differences more specifically in terms of the perceptions of the students via qualitative and quantitative measures is currently underway and the results will be a valuable extension to the results of this study.

Based on empirical estimates of likelihood and priors, both certificate earners and auditors, two groups most engaged with the lectures, showed a higher chance of choosing the Overlay mode suggesting the possibility of this mode being conducive to the viewing characteristics of these learners. The higher chance of a certificate earner choosing the overlay mode over the PiP could be interpreted to mean that improved access to instructor's presence is important to even the most motivated of users of a course in a MOOC environment.

6. LIMITATIONS AND FUTURE WORK

A primary limitation of this study is the lack of a qualitative analysis of user affect and satisfaction with the video mode of their choice. In the absence of the qualitative dimension to our study, most of the quantitative analysis were done based on proxy measures of motivation and navigational intent. Moreover, the measures chosen for the quantitative comparison were approximations based on the course logs with their inherent limitations. A more controlled study encompassing both qualitative aspects and more representative measures of engagement and navigation would shed more light on design guidelines for video lectures.

Our primary measure of engagement, video watching time, only measured the overall interaction with videos without regard to the finer engagement patterns such as the number of pauses and restarts, segments revisited, and playback rate changes that characterize a video view session. Incorporating these details as part of engagement patterns will offer a more refined view of patterns of engagement that are supported by different video presentation styles.

Other aspects for future work in this context would be exploring the preferences based on differences in demographic backgrounds of learners[3]. This would offer key insights about the preferences of a global audience that MOOCs aspire to

[3]Although learner IP address information was available, their potential of being considered as personally identifiable information precluded their inclusion in the analyses.

serve. Another important direction for future work is to explore if the same preferences and outcomes would arise regardless of the demographics the course topic attracts and the immediate functionality of seeing the instructor clearly (i.e content/topic specificity of the course).

7. CONCLUSION

Recognizing the important role that lecture videos play as primary content-bearers of a course in MOOCs, instructional designers are justified in their concerns about the kinds of video presentations that lead to best learning outcomes, keeping video production costs at reasonable levels. In this study we compared two video modes that offered the same set of lectures for a significant duration of a course in programming parallel processors. We found that a significantly large proportion of learners preferred one mode over the other. We hypothesize that the modes primarily differed in their ability to make the instructor's gaze and gestures more directly accessible to learners and that the mode that offered more access to instructor's gestures and eye-gaze was probably the preferred mode by the vast majority of learners. We also hypothesize that these users, possibly owing to the resulting positive affect created by improving the instructor's social presence, showed more engagement with the videos (via larger watch times), preferred the streamed mode of viewing videos (indicating immediacy in user response) and covered a larger proportion of lectures. The results also support the possibility that certificate earners (the most motivated of learners) and auditors (learners who primarily engage with a course by only watching videos) showed a higher chance of choosing the video mode offering better access to instructor's gaze and gestures, suggesting that the mode is perhaps conducive to the viewing characteristics of these learners.

8. REFERENCES

[1] M. Argyle. *Bodily communication*. Routledge, 2013.

[2] S. Bhat and G. Herman. Student perceptions of differences in visual communication mode for an online course in engineering. In *Frontiers in Education Conference, 2013 IEEE*, pages 1471–1473. IEEE, 2013.

[3] J. Borup, R. E. West, and C. R. Graham. Improving online social presence through asynchronous video. *The Internet and Higher Education*, 15(3):195–203, 2012.

[4] L. Breslow, D. E. Pritchard, J. DeBoer, G. S. Stump, A. D. Ho, and D. Seaton. Studying learning in the worldwide classroom: Research into edx?s first mooc. *Research & Practice in Assessment*, 8:13–25, 2013.

[5] M. C. Carlisle. Using you tube to enhance student class preparation in an introductory java course. In *Proceedings of the 41st ACM technical symposium on Computer science education*, pages 470–474. ACM, 2010.

[6] R. C. Clark and R. E. Mayer. *E-learning and the science of instruction: Proven guidelines for consumers and designers of multimedia learning*. John Wiley & Sons, 2011.

[7] G. Cui, B. Lockee, and C. Meng. Building modern online social presence: A review of social presence theory and its instructional design implications for future trends. *Education and information technologies*, 18(4):661–685, 2013.

[8] N. Ford and S. Y. Chen. Matching/mismatching revisited: an empirical study of learning and teaching styles. *British Journal of Educational Technology*, 32(1):5–22, 2001.

[9] P. J. Guo, J. Kim, and R. Rubin. How video production affects student engagement: An empirical study of mooc videos. In *Proceedings of the first ACM conference on Learning@ scale conference*, pages 41–50. ACM, 2014.

[10] P. J. Guo and K. Reinecke. Demographic differences in how students navigate through moocs. In *Proceedings of the first ACM conference on Learning@ scale conference*, pages 21–30. ACM, 2014.

[11] R. F. Kizilcec, K. Papadopoulos, and L. Sritanyaratana. Showing face in video instruction: effects on information retention, visual attention, and affect. In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems*, pages 2095–2102. ACM, 2014.

[12] R. F. Kizilcec, C. Piech, and E. Schneider. Deconstructing disengagement: analyzing learner subpopulations in massive open online courses. In *Proceedings of the third international conference on learning analytics and knowledge*, pages 170–179. ACM, 2013.

[13] J. O. Liegle and T. N. Janicki. The effect of learning styles on the navigation needs of web-based learners. *Computers in Human Behavior*, 22(5):885–898, 2006.

[14] R. E. Mayer. 14 principles based on social cues in multimedia learning: Personalization, voice, image, and embodiment principles. *The Cambridge Handbook of Multimedia Learning*, page 345, 2014.

[15] R. E. Mayer and C. S. DaPra. An embodiment effect in computer-based learning with animated pedagogical agents. *Journal of Experimental Psychology: Applied*, 18(3):239, 2012.

[16] B. Reeves and C. Nass. The media equation: How people respondto computers, television, and new media like real people and places. 2010.

[17] D. T. Seaton, Y. Bergner, I. Chuang, P. Mitros, and D. E. Pritchard. Who does what in a massive open online course? *Communications of the ACM*, 57(4):58–65, 2014.

[18] R. Vertegaal, G. van der Veer, and H. Vons. Effects of gaze on multiparty mediated communication. In *Graphics Interface*, pages 95–102, 2000.

[19] A. Wise, J. Chang, T. Duffy, and R. d. Valle. The effects of teacher social presence on student satisfaction, engagement, and learning. *Journal of Educational Computing Research*, 31(3):247 – 271, 2004.