Chatrooms in MOOCs: All Talk and No Action*

Derrick Coetzee dcoetzee@eecs.berkeley.edu
Armando Fox fox@cs.berkeley.edu
Marti A. Hearst hearst@ischool.berkeley.edu
Björn Hartmann bjoern@berkeley.edu
Computer Science Division and School of Information
University of California, Berkeley, CA 94720, USA

ABSTRACT

We study effects of introducing a real-time chatroom into a massive open online course with several thousand students, supplementing an existing forum. The chatroom was supported by teaching assistants, and generated thousands of lines of discussion by 28% of 681 consenting chat condition participants, mostly on-topic. Despite this, chat activity remained low ($\mu = 8.2$ messages per hour) and we could find no significant effect of chat use on objective or subjective dependent variables such as grades, retention, forum participation, or students' sense of community. Further investigation reveals that only 12% of chat participants have substantive interactions, while the remainder are either passive or have trivial interactions that are unlikely to result in learning.

We also find that pervasive, highly visible chat interfaces are highly effective in encouraging both active and substantive participation in chat. When compared to chat interfaces that are restricted to a single webpage, the pervasive interface exhibits 2.8 times as many users with substantive interactions.

Author Keywords

Massive open online course; MOOC; synchronous; chat; chatroom; retention; participation; experiment.

ACM Classification Keywords

H.5.3. Information Interfaces and Presentation (e.g. HCI): Group and Organization Interfaces; K.3.1. Computers and Education: Computer Uses in Education

INTRODUCTION

Massive open online courses (MOOCs) are online courses which invite large numbers of students (on the order of thousands) to freely enroll. A number of successful large-scale MOOC platforms including edX, Coursera, and Udacity have been developed. In all these platforms, the primary support provided to students who encounter difficulties is through asynchronous threaded forums, which have been called "an essential ingredient of an effective online course" [18].

However, prior work in small-scale online learning suggests that asynchronous mechanisms are most effective when combined with synchronous mechanisms, such as real-time chatrooms (chat) and private messaging [17, 32, 20], and many online courses have effectively incorporated chat [15, 24]. We investigate the question of whether by introducing a chatroom into a MOOC, these results can be extended to the MOOC setting, where there are an order of magnitude more students.

We expected chat to effectively complement established asynchronous forums in MOOCs via a number of mechanisms: it provides a lower barrier to participation, with only minimal steps needed to send messages; it can provide answers in seconds as opposed to the hours typical of forums [4], enabling back-and-forth interactions; and it can encourage community building and forming of relationships, [23] a function for which forums are less suited. A strong sense of community has been identified as important for avoiding attrition, [28] which is a common problem in MOOCs. [3]

Although surveys described our system as "tremendously helpful" and "useful and constructive," in the end we found no significant effects of chat availability on a range of dependent variables including grades, retention, forum participation and sense of community. To understand the disconnect between positive individual reports and the lack of evidence of aggregate effects, we define and analyze *substantive discussions* in MOOC chatrooms, and demonstrate that the proportion of participants meeting this bar is low, suggesting that most students derive no benefit from the chatroom.

Below we discuss prior related work, introduce our chat design and experimental method, summarize results, discuss implications including limitations of the study and possible ways to improve the chat design, and conclude with recommendations for future work.

RELATED WORK

Researchers have investigated the role of synchronous chat in diverse settings, including work environments (usually with small workgroups) [19, 11, 12]; in education settings [15, 2, 17, 14, 27, 20, 32, 24, 26, 6]; in general open settings on Internet Relay Chat (IRC) [21, 1] and around shared video watching in an entertainment context [33].

Here we focus on three relevant aspects of synchronous chat research: the use of chat in online education, user interfaces for chat, and embedded chat interfaces.

L@S'14, March 4–5, 2014, Atlanta, Georgia, USA.
ACM 978-1-4503-2669-8/14/03.
http://dx.doi.org/10.1145/2556325.2566242

Chat in Online Education

Chat has been extensively used in online courses, with numerous works in the learning sciences comparing them to other modes of interaction. Asynchronous mechanisms are found to encourage "in-depth, more thoughtful discussion" [2] while synchronous mechanisms are preferable for "providing a greater sense of presence and generating spontaneity" [13]. In some cases synchronous chat could produce superior learning even compared to face-to-face interaction. [24]

Integrating both produced the best results: they "provide mutual enrichment" [17] in that "chat rooms will enhance and clarify the information that is gathered via asynchronous interactions" [32]. One-on-one synchronous discussion has also been found to "support asynchronous discussions in the formation of a community of inquiry" [20].

However, chats also presented a number of practical hurdles in implementation: "getting students online at the same time, difficulty in moderating larger-scale conversations, lack of reflection time for students" [2]. Although we anticipated that getting participation would be less difficult with an order of magnitude more students, in fact this proves to be a central challenge in the MOOC setting as well.

In reviewing research on text-based community interaction in education, Johnson laments that "[c]omprehensive search of the literature did not result in the identification of a single true experiment [...] random assignment of students to one of two conditions in which one of the conditions is synchronous chat and the other condition is asynchronous discussion." [15] Our study is a true experiment, but we compare synchronous chat against a condition integrating both.

Several studies of chats in online classes focused on synchronous chat *sessions*, which were short, structured, scheduled chats led by instructors [26, 32], sometimes featuring a set of explicit rules or conventions [13]. The chat in our study is unstructured, runs continuously, and is supervised primarily by teaching assistants and other students; we avoid implementing explicit policies in order to investigate spontaneous usage. Although both types of chats have advantages, in our MOOC setting a continuously running chat is able to complement the existing forum by providing more rapid responses, and avoids unfairly excluding students based on time zone.

Chat Interfaces

HCI researchers have attempted to overcome known problems of synchronous chat — including overlapping conversations, difficulty following conversation threads, and poor conveyance of tone and emotion — by designing alternative chat interfaces like comics [16], temporal message flows [31], conversation trees [25] or automatically clustered groups [30]. To simplify implementation, we use a web-based chatroom with a simple, conventional interface. Users can learn strategies for repairing misunderstandings [19] and for conveying tone and emotion [10], partly mitigating the issues outlined above.

A number of *persistent chat systems* record chat information in such a way that it can be used later. Among these, our system is most similar to Babble, [8] in that it stores a log of conversations and allows students and staff to access it at any time by scrolling up in the interface. Unlike BackTalk, [9] which relies on user annotations to transform chat data into structured persistent data for review, we simply allow students to use the asynchronous forum to persist discussions in a structured manner as needed.

Embedded Chat Interfaces

One of our primary contributions is the investigation of embedded chat interfaces, placed on the same page next to video lectures, assignments, and quizzes. Cummings and Guerlain investigated embedding chat into a military system in order to enable "secondary tasking" (responding to instructions and queries) using "spare mental capacity" [5]. They found that chat activity generally degraded during demanding primary tasks (e.g. missile retargeting), but that some operators instead fixated on the chat resulting in lower performance on the primary task. Although this raises the possibility that the embedded chat may damage course outcomes by disrupting private study, we failed to identify significant negative effects.

Work on dialogue in collaborative learning systems distinguishes between "parallel tools," which "do not assure any coordination between the discourse and disciplinary representations," and "embedded tools" such as annotation tools which embed comments directly into the artifact under discussion. [7] In this vocabulary, our embedded chat is technically a parallel tool, since the chat is in proximity to but separate from learning artifacts, and chat users can and do discuss unrelated topics. Although properly embedded tools provide greater context for communication, a conventional chatroom is able to preserve chronological order of discussion, avoiding problems in which "the record of discourse is fragmented across the artifact."

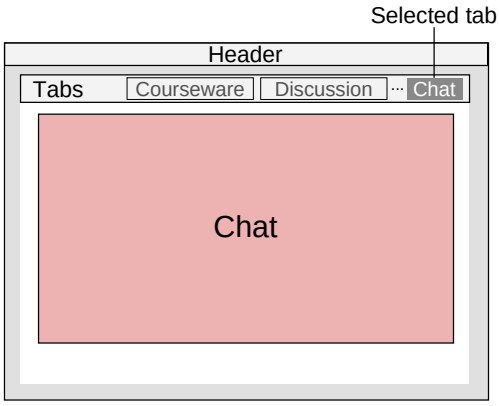
METHOD

We conducted a between-subjects field experiment on a seven-week, open-enrollment software engineering course offered on the edX platform ("CS169.1x: Software as a Service" from the University of California, Berkeley).

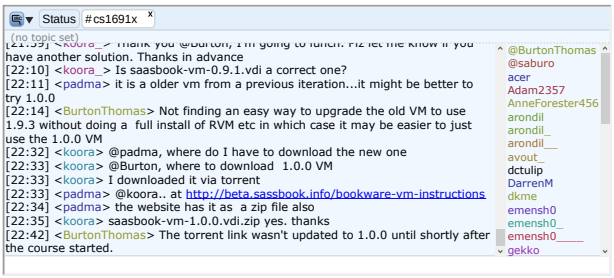
We explored two different methods of integrating chat functionality into the site: **Chat tab:** A prominent "Chat" tab is added to a list of links at the top of the site. This link takes students to a dedicated page where they can participate in the chat (Figure 1a). **Embedded chat:** In addition to the chat tab, every page of the site, including lectures and assignments, has a panel embedded where the live chat is displayed (see Figure 1c). Both methods display the same shared chatroom, and both methods display chat history automatically upon joining, enabling students to examine past messages. Figure 1 compares the two methods of integration. The chat interface is conventional, with messages at left and a user list at right (see Figure 1b). The field experiment also had a **Control** condition in which students who consented to participate in the study were shown an unmodified edX interface with no chat.

Participants

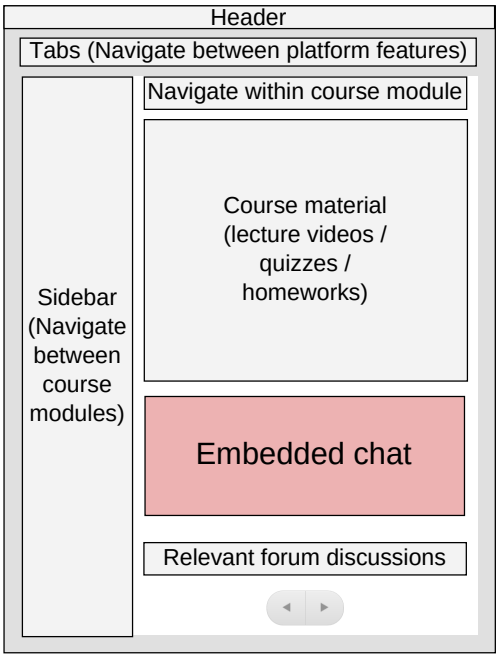
14381 students were enrolled in the course as of January 2014. Of these, 1344 (9.3%) consented to participate in our



(a) Chat tab interface, accessible to both the "chat tab" and "embedded chat" groups by clicking the "Chat" button in the upper right. The chat panel is wider and taller than the embedded chat, filling most of the browser window.



(b) Web chat interface, based on `qwebirc`. Messages with times are on the left, users are on the right, and messages are entered at bottom. Names have been changed.



(c) Embedded chat interface, accessible only to the "embedded chat" group. A smaller chat panel is presented below the lecture, quiz, or homework the student is currently interacting with. In some cases, the student may need to scroll to see it. If the student is not in the "embedded chat" group, the chat panel is hidden but the page is otherwise identical.

Figure 1: The chat interface is presented in two different ways: on a dedicated page (*top left*) and embedded in a smaller form underneath the course materials (*right*). The detailed layout of the chat interface is also shown, including an actual conversation (*bottom left*).

experiment. 509 students were assigned to the *no chat* condition, 409 to the *chat tab* condition, and 426 to the *embedded chat* condition. The variation in these counts is due to how students were assigned to groups, by applying a hash function to their username. These numbers suggest that a MOOC study seeking a specific number of subjects should target a course with about 10 times as many enrolled students, due to rapid attrition during the earliest phase of the course.

Hypotheses

Our three experimental groups enable two types of controlled comparisons: the comparison of users with and without access to chat ("chat tab" and "embedded chat" groups combined versus control group), and the comparison of users with and without embedded chat integration ("chat tab" versus "embedded chat" group). Direct comparisons cannot be made between the control group and either the chat tab or embedded chat group on their own, because these two groups use the same chatroom and so influence one another's behavior.

Our primary question is to whether access to chat provides an objective advantage in learning as measured by course outcomes such as grades and retention/attrition (the duration the student remains in the course before dropping). A secondary goal is to establish whether chat exhibits the predicted advantages over the asynchronous forum, promoting a sense of community and lowering the bar to active participation. Finally, we investigate if the higher visibility of the embedded chat as compared to the chat tab design encourages more active participation. We focus on active participation in the chat (sending messages), because we lack any means to measure passive participation (reading messages).

Course Outcomes

It is unclear whether chat should be expected to benefit or hurt outcomes: on one hand it supports thoughtful discussion on course material and sense of community, while on the other it may distract from private study. We anticipate that the advantages will outweigh the disadvantages.

H1 Students in the chat conditions have higher retention than non-chat students.

H2 Students in the chat conditions have higher course grades than non-chat students.

Comparing Chat and Forum Activity

H3 The proportion of active chat users (among users with access to chat) is greater than the proportion of active forum users (among all study participants).

H4 Chat availability may decrease the number of forum posts by diverting students from the forum.

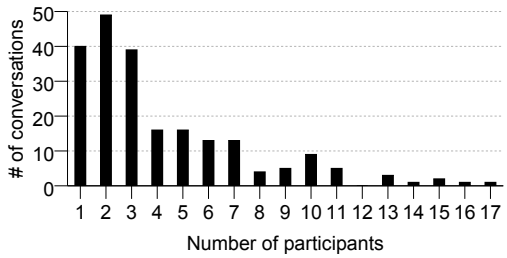


Figure 2: When the chat is divided into contiguous conversations, 59% had at most 3 participants, while the rest had a larger number, up to 17.

Sense of Community

Rovai's Classroom Community Scale [22] is a survey device based on 20 Likert-scale questions, such as "I feel that it is hard to get help when I have a question." It is designed to measure a student's subjective sense of being part of a community in the context of a course. Since interacting with others promotes community, particularly in a social setting, we anticipate that:

H5 Students in the chat conditions have higher sense of community scores than no-chat students.

Differences between Embedded and Chat Tab Activity

We hypothesize that the greater visibility of the embedded chat interface will lead to increased participation.

H6 More students in the *embedded chat* condition will post to the chat than in the *chat tab* condition (because of the visibility of the chat interface).

H7 Students who post in the *embedded chat* condition will post more messages than students in the *chat tab* condition (again, because of visibility.)

RESULTS

Over the duration of the course, 8980 messages were posted. 2169 messages (24.2%) were posted by administrative users (teaching assistants and other course staff); 6811 (75.8%) were posted by students.

Chat conversations

Separating overlapping conversations in chat data is challenging, and there are many techniques for doing so [29]. However, in our case overlapping conversations were rare due to low chat activity ($\mu = 8.2$ messages per hour), permitting trivial segmentation of conversations based on a pause between conversations of at least 1 hour. With this segmentation method, there were 216 conversations with a median length of 11.5 messages and a median of 3 and mode of 2 participants (see Figure 2). Most conversations were short, with 49% of 10 messages or less. In 40 or 18.5% of conversations, a single user spoke and no one responded.

Activity occurred throughout different days of the week, with a notable spike on Wednesdays, and a dip Fridays, GMT (see Figure 3). The spike in activity is likely related to course deadlines: assignments and quizzes were due on Wednesdays

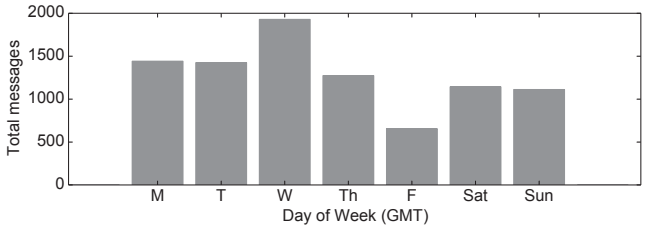


Figure 3: The histogram of messages by day of week shows high activity on Wednesdays, corresponding with course deadlines; and low activity on Fridays.

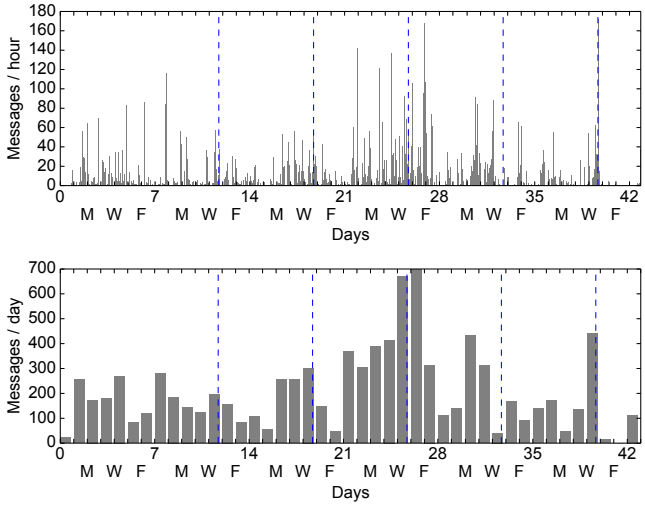


Figure 4: Chats were very irregular and bursty. Top: Histogram of chat messages sent per hour. Bottom: Histogram of chat messages per day. Dashed blue lines indicate homework and quiz due dates.

17:00 GMT, and new lectures and assignments were released on Thursdays at 12:00 GMT.

Chat conversations often appeared in bursts throughout a day — see, e.g., the large spikes in hourly activity on days 7 and 26 in Figure 4, Top. Again, chat activity is sometimes clearly correlated with homework and quiz deadlines on Wednesday afternoons, e.g., on days 18 and 39 (see dashed lines indicating course milestones in Figure 4, Bottom). However, this is not always the case: there is a high volume of chat messages throughout week four; and no spike in messages around the penultimate assignment on day 32.

When aggregating messages by time of day across the entire corpus, the chat exhibits a pattern of decreased activity during night-time hours, GMT (see Figure 5). However, due to wide distribution of students across time zones, these are difficult to interpret. This irregular activity pattern is likely unique to MOOCs that draw students from across the globe and is an important difference to traditional online courses with geographically limited audiences.

Course Outcomes for Chat Users vs. Non-Chat Users

We determined how long each student (excluding staff) remained in the course, based on the time of their last interac-

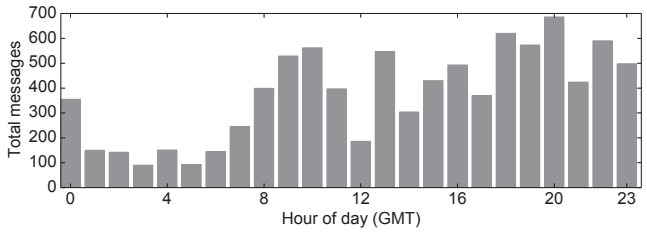


Figure 5: The histogram of cumulative chat messages by hour of day shows multiple periods of high and low activity.

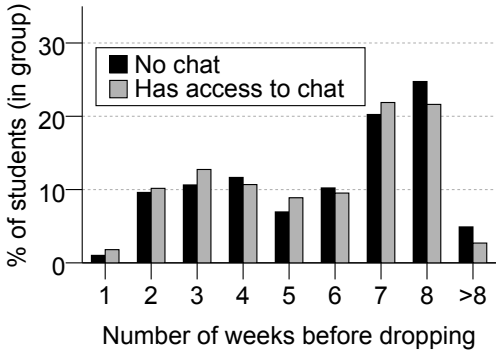


Figure 6: Histogram of how many weeks students with and without access to the chatroom spent before ceasing interaction with the course website. Drop rates for both groups remain similar for both groups at each point in the course. No significant difference between them was found ($p > 0.06$).

tion with any element of the course website. We compared retention times of the control group (non-chat) against the "chat tab" and "embedded chat" groups combined (chat students). Medians were 36.8 and 35.9 days, respectively, and no significant difference could be shown (Mann-Whitney U = 137313.0, $n = 418,694$, $p > 0.06$). This is reflected in Figure 6, which shows similar drop rates of the two groups at each point in the course. Hypothesis H1 is not supported.

Because different students drop at different points, the most appropriate way to compare grades is by comparing grades on particular quizzes or assignments, and restricting the analysis to students who completed the quiz or assignment. We compared score distributions using two-sample Kolmogorov-Smirnov tests, which did not find significant differences for any assignment (see Figure 7). Hence, H2 is not supported.

Comparing Chat and Forum Participation

To determine whether access to chat affects forum use, we compared the proportion of active forum users (who posted at least one post of any type) in the control group (non-chat) and in the "chat tab" and "embedded chat" groups combined (chat students). We found 118 (23%) of the 509 non-chat users posted in the forum, while 201 (24%) of 835 chat users posted in the forum. Fisher's test finds no significant difference ($p > 0.7$). Hypothesis H4 is not supported.

Summing these, 319 (24%) of 1344 study participants posted on the forum, while 191 (23%) of 835 chat users sent at least

Quiz	Median (non-chat)	Median (chat)	Max score	n_1	n_2	D	p
0	12.5	13	13	270	223	0.03	> 0.9
1	9	8	12	188	159	0.09	> 0.5
2	12	12	16	124	119	0.08	> 0.8
3	9	9	11	107	97	0.09	> 0.7

HW	Median (non-chat)	Median (chat)	Max score	n_1	n_2	D	p
0	300	300	300	247	209	0.07	> 0.6
1	400	400	400	190	163	0.06	> 0.8
1.5	400	400	400	137	136	0.04	> 0.9
2	93	93	100	107	99	0.06	> 0.9
3	500	500	500	94	82	0.09	> 0.8
4	500	400	500	70	62	0.12	> 0.7

Figure 7: Comparison of median grades on assignments (HW=homework). Medians were similar for all of them, and the two-sample Kolmogorov-Smirnov test did not identify a significant difference in the grade distribution for any of them.

one message. Fisher's test finds no significant difference between the proportion of active forum and chat users ($p > 0.6$), and Hypothesis H3 is not supported. This also calls into question the original assumption that the chat will lower the bar to participation compared to the forum. Although the percentages are very similar, it is not the same users using both systems, as shown in Figure 8.

Participation in Chat Tab and Embedded Conditions

More students in the *embedded chat* condition posted to the chat than in the *chat tab* condition: 54 of 399 users (13.5%) in the *chat tab* condition participated actively (posted at least one message), while 128 of 419 users (30.5%) in the *embedded chat* condition were active participants. This difference was statistically significant (Fisher's exact test, $p < 0.0001$). Hypothesis H6 is supported.

As in many online communities, the number of messages sent by users were characterized by a long-tailed distribution in which a few users post very frequently, but most users post very few messages. Of the active participants, students in the *chat tab* condition posted a median of 3.5 messages, while students in the *embedded chat* condition posted a median of 4 messages (see Figure 9). Figure 9, Bottom also shows that users at comparable percentiles tend to post more messages in the *embedded* condition than in the *chat tab* condition for most of the distribution. However, a two-sample Kolmogorov-Smirnov test ($D = 0.18$, $p = 0.18$) showed that the difference in distributions was not statistically significant. Our results are thus inconclusive whether an embedded chat interface leads students who have already decided to participate in the chat to participate more. Hypothesis H7 is thus not supported.

Survey Results

We offered two optional surveys in the course, a pre-survey at the time the chat was deployed, and another survey after day 25 to gather retrospective information about the chat. (Administering surveys at the very end of a MOOC is ineffective

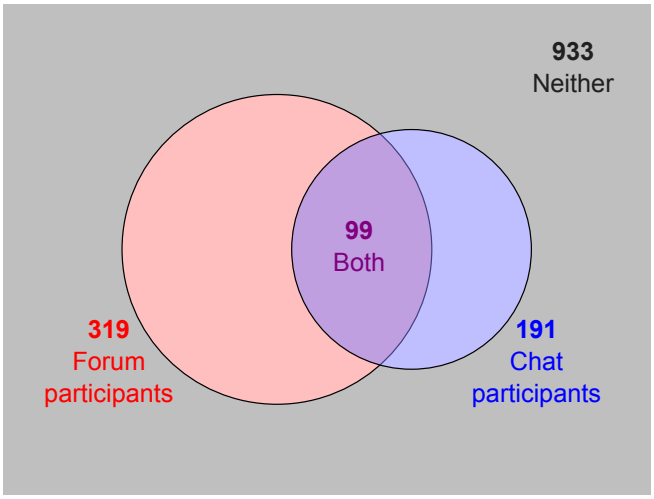


Figure 8: Venn diagram of communication modes (to scale). 319 posted on the forum at least once, 191 participated in the chat at least once, and 99 participated in both. In particular, 92 (48%) of the active chat participants never posted on the forum, including 4 of the most active 20 students in the chat. 933 (69%) of the 1344 study participants used neither mode of communication. This suggests that the chat is able to effectively involve some students who would not otherwise interact with others, but most students still participate only passively. Note that the forum was also used by students not participating in the study; they are excluded above.

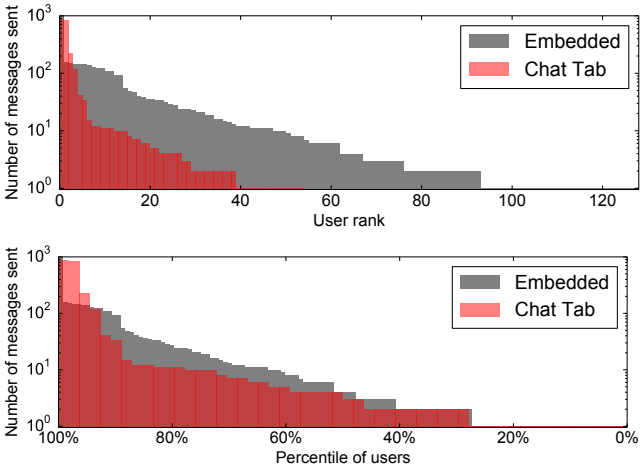


Figure 9: Rank-order plot of messages sent by users in the chat tab and embedded conditions. Top: Absolute user ranks shows higher participation in the embedded condition. Bottom: Percentile plot shows higher participation of the middle of the distribution for the embedded condition; however, this difference is not statistically significant.

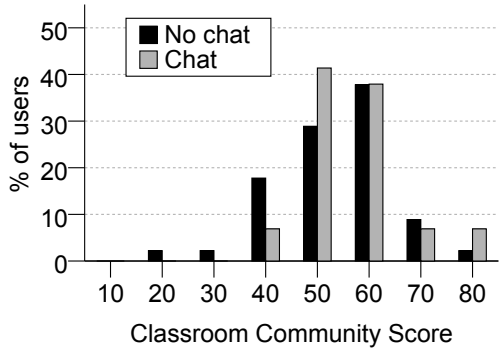


Figure 10: Comparison of sense of community scores for non-chat (control group) and chat (other groups) students. Medians were near-identical (50 and 51) and no significant difference was found ($p > 0.2$).

as most students will have dropped the course by then [4].) The pre-survey was given to all students in the course and had 1486 responses, while the later survey was offered to all subjects in the chat experiment including the control group and had 112 responses. The latter represents 48 (9.2%) of 519 students in *no chat*, 32 (7.8%) of 409 students in *chat tab*, and 32 (7.5%) of 426 students in *embedded chat*.

In the pre-survey, despite the fact that the course was targeted at software developers, we found that 45% of students had no prior experience with chatrooms, and only 6% used them frequently. This inexperience, combined with our system's lack of training or tutorials, may be another factor underlying low participation.

We applied Rovai's Classroom Community Scale [22] to measure the subjective sense of community experienced by all study groups. The median Rovai sense of community scores for the no-chat and chat groups were 50 and 51, respectively, and no significant difference could be found (Mann-Whitney U=1212.5, $n = 45, 58$, $p > 0.2$). Hypothesis H5 is not supported. Figure 10 compares the distributions.

Passive participation in the chat was reported more often in the embedded condition (81% of students read the chat at some point) than in the chat tab condition (64%), but this difference is not significant (Fisher's exact test, $p > 0.2$). In the embedded chat condition, all surveyed students were aware of the chat, whereas only 86% were aware of it in the chat tab condition, and this difference was marginally significant ($p < 0.05$). On the other hand, in the embedded chat group, 48% found the embedded chat "distracting or annoying."

Most students (16 out of the 18 students who answered the question) reported that TAs and students were equally helpful in the chat; this is consistent with the data, showing a mixture of students and TAs among the most prolific chatters. "Answering specific questions about course content" was the most common purpose for which the chatroom was used (20 of 25 who reported using the chatroom reported using it for this purpose), and among the 14 respondents who responded to others, an altrustic desire to help others was the most reported reason (13 of 14). 72% reported they got a useful response from others either "sometimes" or more often.

Anecdotally, a number of students surveyed reported strongly positive experiences with the chatroom: "I find the chat to be tremendously helpful in the clarification of homework problems. [...] Due to my schedule, I was often down to the wire for several submissions, and a chatroom allowed for much faster responses than something like an emailed question." "It was great to get instant feedback, quick answers, and encouragement."

Students reported using the chatroom and forum in combination: "Sometimes, I post the discussion & give the link in chat room." Multiple students also reported positive passive experiences: "Many useful and constructive real time conversations on topics even though I wasn't actively participating." "What other people ask in the chat room is also useful."

Some students felt the chat was unhelpful for them personally: "I feel that most of the students are below my experience in IT, so did not feel any need to chat." "[M]ost of the comments in the chatrooms related to future course [material] as opposed to what I was working on." Others struggled with its unstructured nature, preferring the more structured forum: "[T]he whole chatting stuff looks too unstructured." "I found course forum more helpful for me, mainly because information there has some structure applied."

Classifying Users by Level of Interaction

The above results present a paradox: the chat is shown to engage a number of users who otherwise have limited participation, and produces great anecdotal experiences, yet we can detect no effect of chat availability on any objective dependent variables, including grades, retention, forum participation, and subjective sense of community.

One explanation for this is that benefits of the chat accrue only to a relatively small number of active users. When we compare subjects who sent at least one message in chat (active in chat) to subjects who did not, we find that they remained in the course a median of 7.2 days longer (45.1 vs. 37.9), a significant difference (Mann-Whitney U = 80795.0, $n = 182, 1084$, $p < 0.0001$). This is weak evidence, since active users are self-selected and more dedicated students are more likely to participate in chat, but it suggests a possible explanation.

Even among active users, many had limited interaction with the chat. To quantify this, each student was manually classified into one of the following mutually exclusive groups:

- **Tester**: Sent only test messages; no meaningful content.
- **Greeter**: Sent only messages containing greetings.
- **Socializer**: Only discussed off-topic or irrelevant material.
- **No response**: Asked a question but received no response.
- **No acknowledge**: Asked a question and received a response but showed no sign of noticing the response.
- **Acknowledged**: Asked a question, received a response, and clearly acknowledged the response.
- **Answerer**: Student is not in **Acknowledged** category, but did respond to questions of others.

Davidson-Shivers et al. similarly coded individual chat messages as substantive or non-substantive according to a set of nine categories, some of which align with ours (e.g. Responding is related to our Answerer role, Chatting is related to our Socializer role, and Uncodable is related to our Tester role) [6]; however, we categorize users rather than individual messages. For example, a student who both greets and responds to questions would be in the **Answerer** category; a student who both acknowledges responses to questions and answers questions would be in the **Acknowledged** category.

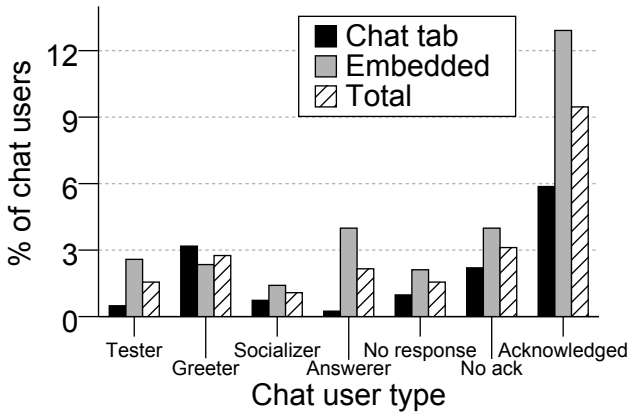


Figure 11: Comparison of sizes of 7 types of chat participants (see text). Only the Acknowledged and possibly the Answerer types represent substantive participation, which here comprise at most 12% of the chat students as a whole. The embedded chat condition has a strong advantage over the chat tab condition in the substantive categories ($p < 0.001$ for Acknowledged and Answerer). The remaining 78% of chat users not shown had no participation in the chat whatsoever.

The **Acknowledged** group above is meant to capture our understanding of a minimum bar for substantive chat usage that can produce learning. Although students in the **No acknowledge** category receive responses, we assume they did not notice them; this category features a median time of almost 7 minutes between question and response, suggesting that questioners may have diverted attention to other tasks before their response arrives. Figure 11 summarizes the size of these categories over our user base.

Although **Acknowledged** is the largest category, representing 41% of all active chat users, it is still only 9.5% of all students with access to chat, and even including **Answerer** only raises that to 12%. When we compare retention of students in these two categories to active chat participants in the other categories, we find a median difference of 4.7 days (47.3 vs. 42.5 days), a significant difference (Mann-Whitney U = 3254.0, $n = 98, 84$, $p < 0.01$). When we compare retention of students in these two categories to retention of all other subjects, the median difference is 9.2 days (47.3 vs 38.1 days), and this is significant (Mann-Whitney U = 41285, $n = 98, 1168$, $p < 0.0001$). Although again these are weak results due to self-selection, higher retention is clearly correlated with more substantive participation.

With 88% of chat students failing to engage in substantive participation, even strong improvements by the few users who do engage in it could not substantially shift the outcomes of the group as a whole. Larger sample sizes would be required to reliably detect such a small change.

We found earlier that the embedded chat group is about twice as active as the chat tab group but Figure 11 shows an even stronger advantage in user categories with substantive inter-

actions (**Acknowledged** and **Answerer**). Fisher's test shows that the embedded chat group is significantly higher in both ($p < 0.001$ for both). Overall, about 17% of embedded chat users had substantive interactions compared to only 6% of chat tab users, or 2.8 times as many.

In some circles chatrooms are controversial because they are "viewed as recreational, as opposed to educational" [15]. In this context, our data repudiates this idea, with users focused on off-topic discussion being the rarest of all types. This may be explained by multiple factors, including continual monitoring by TAs, demographics of the course, and so on.

DISCUSSION

Recommendations for instructors

Based on our findings, we can make three specific recommendations to instructors interested in using chat in MOOCs:

- **Should I use chat?** Chat is safe to use; there is no evidence that it degrades learning or forum participation. Strong praise from surveyed students, as well as evidence that it can engage some students who don't participate in the forum, suggests that it may be worthwhile, but the operating cost must be kept low to justify a system that benefits only a small number of students.
- **How should chat be integrated into my course website?** Pervasive, highly-visible interfaces are the best choice for maximizing substantive participation in chat. Although some students found them annoying, we found no evidence that they adversely impact learning.
- **How do I ensure chats remain on-topic?** In our setting, volunteer teaching assistants moderated the chatroom on an *ad hoc* basis. This was enough to ensure that most active students engaged in substantive, on-topic conversations.

Limitations due to low chat activity

The low proportion of substantive users of chat makes it unexpectedly difficult to reach critical mass for an effective chatroom, even in relatively large courses. In our study, the consent process favored students with a pre-existing interest in chatrooms, and yet only 12% of users had substantive interactions; in a setting without a consent hurdle, this percentage may drop even further. Users with continued engagement over time were even rarer. The result was a system that was anecdotally valuable for a few students, yet unused by most.

One natural strategy is to attempt to extend the positive experience that some chat participants anecdotally received to a larger student population by increasing the number of students with substantive conversations in chat. Although we can't predict whether a such a higher-activity chat would objectively benefit learning outcomes, it forms a useful starting place for refining our design. Following are a few strategies that might be used to increase activity:

- Increasing the total number of students. Linear extrapolation from our results suggests that with 4600 students we might see about one message per minute, resulting in more regular/constant chat activity.
- Increasing the number of students using chat, perhaps with more aggressive UI cues. Indeed, 15% of users with the chat tab interface were unaware of the chat's existence, a problem easily corrected with a more pervasive interface. Some survey respondents requested omnipresent interfaces such as chat overlays permanently pinned in the corner of every page with notifications, as used today on Facebook and Google+. However, an overly aggressive UI can also become more distracting and annoying for students.
- Restricting chat availability to certain hours to increase chat density. However, this limits chat's usefulness for getting timely answers to urgent questions, and may unfairly disadvantage students in certain time zones.
- Increasing the percentage of users with substantive interaction. While Figure 11 suggests that participation can be improved by adding sufficient helpers to address all student concerns, the same figure suggests that we can expect a gain of at most 10 percentage points relative to the conditions in this experiment. Most of the nonparticipation occurs among the 78% of chat users who never actively interact with the chat at all.
- Addressing conditions reported in survey that caused some students not to participate in chat, such as being too busy with other coursework or being too far behind to contribute to or benefit from technical chat conversations. Separate chatrooms for different parts of the course might mitigate this problem.
- Addressing cultural or personal factors that may inhibit students from using chat. For example, some may be shy or feel reluctant to ask questions or offer responses that might make them seem ignorant; some may be unwilling to chat under their edX username, which we enforced; some may be uncomfortable due to poor command of English. Further investigations into and designs to accommodate such traits could be valuable.

While there are many strategies to improve activity, an open question is whether chat can deliver the hypothesized benefits of synchronous interaction. While we failed to find evidence of these benefits, it would be hasty to claim that they do not exist. The combination of relatively small sample sizes, low participation, and assessments with poor discrimination (see next section) implies that even a strong benefit among active chat participants could have evaded detection. Further work with larger student populations, more aggressive interfaces, and more challenging assessments may be able to uncover benefits that we could not.

Data limitations

In addition to the small percentage of subjects affected by the chat, other factors limit our ability to detect differences between study groups. The assessments used in the course under study had high median scores, often perfect scores, implying that they're not challenging enough to distinguish students with an average understanding of the topic from highly competent ones. This is partly because the course under study permitted resubmissions, and currently only data for the final

submission is available. Because the students who were most active in chat were often already highly motivated, improvements in their understanding may not be measurable through these assessments.

Due to incomplete information about the focus of the student's attention, we could not reliably determine when the user was reading the chat interface. This is particularly true for the embedded chat group, who always had the chat open whenever visiting the course site, but rarely looked at it. Although comments in the surveys suggest users may have benefitted from passive participation, this cannot be measured or tested. Similarly, we could not determine whether or not students read or acted upon a response unless they explicitly acknowledged it. To some extent this may be addressed with additional software support (e.g. tracking control focus and browser window position), but we still cannot eliminate the possibility that the student's attention is directed to another part of the screen.

Complementary and competitive technologies

During the course, 340 web links were posted to the chat, both to respond to questions and to help coordinate further interaction. Threads on the course forum were linked 24 times. The *pastebin* website was used to share large code samples on 11 occasions. The course involved both live video tutorials and pair programming exercises with screensharing, and both teaching assistants and students used the chatroom to recruit for these activities. Linking relevant web resources to answer questions was also common. These resources all played a complementary role to the chat, providing essential features that the chat is not intended to, without subsuming its function entirely.

On the other hand, some discussions that might have been useful in the chat were moved out of the chat because of its limited functionality. One discussion was moved to Google Hangouts in order to use screensharing; another was moved to Skype chat because of better Unicode support. When discussions leave the chat, the chatroom loses active users, passive users can no longer benefit from the discussion, and course staff cannot record and monitor them. This is an instance of a more general problem for MOOC research: it's difficult to capture all the data associated with student behavior or engagement in a MOOC.

CONCLUSION AND FUTURE WORK

In this work, we introduced a chatroom into a MOOC. Although we found no significant effect on dependent variables such as grades, retention, forum participation, and sense of community, this is unsurprising given that only 12% of students with access to chat engaged in substantive chat interactions. We found that the use of pervasive, highly visible interfaces increased substantive interactions by 2.8 times as compared to interfaces contained to a single page.

In future work, an important direction is exploring design changes that both increase the number of observers in the chat, and increase the percentage of users with substantive interactions. These include: the use of more pervasive interfaces which are always on-screen in a consistent location, the use of notification features to indicate when a user is mentioned or when the chat is most useful to them, the ability to chat under a different pseudonym, restricting chats to particular time periods, and so on.

Another possibility is tighter integration of the chat and forum: currently chat users link to relevant forum threads that already exist, but knowledge is rarely transferred in the other direction, from the chat to the forum. Questions which cannot be answered in a timely manner could be automatically transferred to the forum so that they aren't displaced by new messages and lost. Conversely, new or outstanding messages on the forum could announce themselves on the chat in order to decrease forum response time.

Finally, just as reputation systems help in forums to decrease response time, identify reliable actors, and enforce community norms, a similar system could serve the same purpose in a chatroom. By providing points in exchange for helpful questions and responses, such a system could effectively "bootstrap" chat communities which have just been created for a course and have no established community norms to draw upon.

ACKNOWLEDGMENTS

We thank the World TAs of CS169.1x for coordinating support of students in the chatrooms. This work was partially funded by the National Science Foundation under award IIS-1149799 and the National Endowment for the Humanities under grant HK-50011.

REFERENCES

1. Bechar-Israeli, H. From <Bonehead> to <cLoNehEAd>: Nicknames, Play and Identity on Internet Relay Chat. *Journal of Computer-Mediated Communication 1*, 2 (1995), 00.
2. Branon, R., and Essex, C. Synchronous and asynchronous communication tools in distance education. *TechTrends 45*, 1 (2001), 36–36.
3. Clow, D. MOOCs and the funnel of participation. In *LAK '13: 3rd International Conference on Learning Analytics & Knowledge* (2013).
4. Coetzee, D., Fox, A., Hearst, M. A., and Hartmann, B. Should your MOOC forum use a reputation system? In *Proceedings of the 2014 Conference on Computer-Supported Cooperative Work*, ACM (New York, NY, USA, 2014).
5. Cummings, M., and Guerlain, S. Using a chat interface as an embedded secondary tasking tool. *Human performance, situation awareness and automation: Current research and trends. HPSAA II 1* (2004), 240–248.
6. Davidson-Shivers, G. V., Muilenburg, L. Y., and Tanner, E. J. How do students participate in synchronous and asynchronous online discussions? *J. EDUCATIONAL COMPUTING RESEARCH 25*, 4 (2001), 351–366.
7. Dimitracopoulou, A. Designing collaborative learning systems: current trends & future research agenda. In

Proceedings of the 2005 conference on Computer support for collaborative learning: learning 2005: the next 10 years!, CSCL '05, International Society of the Learning Sciences (2005), 115–124.

8. Erickson, T., Smith, D. N., Kellogg, W. A., Laff, M., Richards, J. T., and Bradner, E. Socially translucent systems: social proxies, persistent conversation, and the design of "babble". In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, CHI '99, ACM (New York, NY, USA, 1999), 72–79.
9. Fono, D., and Baecker, R. Structuring and supporting persistent chat conversations. In *Proceedings of the 2006 20th anniversary conference on Computer supported cooperative work*, CSCW '06, ACM (New York, NY, USA, 2006), 455–458.
10. Gajadhar, J., and Green, J. An analysis of nonverbal communication in an online chat group. *The Open Polytechnic of New Zealand, Working Paper 23* (Mar. 2003).
11. Handel, M., and Herbsleb, J. D. What is chat doing in the workplace? In *Proceedings of the 2002 ACM conference on Computer supported cooperative work*, CSCW '02, ACM (New York, NY, USA, 2002), 110.
12. Herbsleb, J. D., Atkins, D. L., Boyer, D. G., Handel, M., and Finholt, T. A. Introducing instant messaging and chat in the workplace. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '02, ACM (New York, NY, USA, 2002), 171178.
13. Hines, R. A., and Pearl, C. E. Increasing interaction in web-based instruction: Using synchronous chats and asynchronus discussions. *Rural Special Education Quarterly 23*, 2 (Mar. 2004), 33.
14. Ingram, A. L., Hathorn, L. G., and Evans, A. Beyond chat on the internet. *Computers & Education 35*, 1 (2000), 21 – 35.
15. Johnson, G. Synchronous and asynchronous text-based CMC in educational contexts: A review of recent research. *TechTrends 50*, 4 (2006), 46–53.
16. Kurlander, D., Skelly, T., and Salesin, D. Comic chat. In *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques*, SIGGRAPH '96, ACM (New York, NY, USA, 1996), 225236.
17. Ligorio, M. B. Integrating communication formats: synchronous versus asynchronous and text-based versus visual. *Computers & Education 37*, 2 (2001), 103 – 125.
18. Mak, S., Williams, R., and Mackness, J. Blogs and forums as communication and learning tools in a MOOC. In *International Conference on Networked Learning 2010*. 275–285.
19. O'Neill, J., and Martin, D. Text chat in action. In *Proceedings of the 2003 international ACM SIGGROUP conference on Supporting group work*, GROUP '03, ACM (New York, NY, USA, 2003), 4049.
20. Oztok, M., Zingaro, D., Brett, C., and Hewitt, J. Exploring asynchronous and synchronous tool use in online courses. *Computers & Education 60*, 1 (2013), 87 – 94.
21. Rintel, E. S., and Pittam, J. Strangers in a strange land interaction management on internet relay chat. *Human Communication Research 23*, 4 (1997), 507534.
22. Rovai, A. Development of an instrument to measure classroom community. *The Internet and Higher Education 5*, 3 (FebMar 2002), 197–211.
23. Rovai, A. P. Building classroom community at a distance: A case study. *Educational Technology Research and Development 49*, 4 (2001), 33–48.
24. Schoenfeld-Tacher, R., Mcconnell, S., and Graham, M. Do No HarmA Comparison of the Effects of On-Line vs. Traditional Delivery Media on a Science Course. *Journal of Science Education and Technology 10*, 3 (2001), 257–265.
25. Smith, M., Cadiz, J. J., and Burkhalter, B. Conversation trees and threaded chats. In *Proceedings of the 2000 ACM conference on Computer supported cooperative work*, CSCW '00, ACM (New York, NY, USA, 2000), 97–105.
26. Spencer, D. H., and Hiltz, S. A field study of use of synchronous chat in online courses. In *System Sciences, 2003. Proceedings of the 36th Annual Hawaii International Conference on* (2003), 10 pp.–.
27. Stein, D. S., Wanstreet, C. E., Glazer, H. R., Engle, C. L., Harris, R. A., Johnston, S. M., Simons, M. R., and Trinko, L. A. Creating shared understanding through chats in a community of inquiry. *The Internet and Higher Education 10*, 2 (2007), 103 – 115.
28. Tinto, V. *Leaving College: Rethinking the Causes and Cures of Student Attrition*. University of Chicago Press, 1993.
29. Uthus, D. C., and Aha, D. W. Multiparticipant chat analysis: A survey. *Artificial Intelligence 199200*, 0 (2013), 106 – 121.
30. Vigas, F. B., and Donath, J. S. Chat circles. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, CHI '99, ACM (New York, NY, USA, 1999), 916.
31. Vronay, D., Smith, M., and Drucker, S. Alternative interfaces for chat. In *Proceedings of the 12th annual ACM symposium on User interface software and technology*, UIST '99, ACM (New York, NY, USA, 1999), 1926.
32. Wang, A., and Newlin, M. Online Lectures: Benefits for the virtual classroom. *T.H.E. Journal 29*, 1 (2001), 17–24.
33. Weisz, J. D., Kiesler, S., Zhang, H., Ren, Y., Kraut, R. E., and Konstan, J. A. Watching together: integrating text chat with video. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '07, ACM (New York, NY, USA, 2007), 877886.