Monitoring MOOCs: Which Information Sources Do Instructors Value?

Kristin Stephens-Martinez
EECS Department
UC Berkeley
Berkeley, CA 94720
ksteph@cs.berkeley.edu

Marti A. Hearst
School of Information
UC Berkeley
Berkeley, CA 94720
hearst@berkeley.edu

Armando Fox
EECS Department
UC Berkeley
Berkeley, CA 94720
fox@cs.berkeley.edu

ABSTRACT
For an instructor who is teaching a massive open online course (MOOC), what is the best way to understand their class? What is the best way to view how the students are interacting with the content while the course is running? To help prepare for the next iteration, how should the course's data be best analyzed after the fact? How do these instructional monitoring needs differ between online courses with tens of thousands of students and courses with only tens? This paper reports the results of a survey of 92 MOOC instructors who answered questions about which information they find useful in their course, with the end goal of creating an information display for MOOC instructors.

The main findings are: (i) quantitative data sources such as grades, although useful, are not sufficient; understanding the activity in discussion forums and student surveys was rated useful for all use cases by a large majority of respondents, (ii) chat logs were not seen as useful, (iii) for the most part, the same sources of information were seen as useful as found in surveys of smaller online courses, (iv) mockups of existing and novel visualization techniques were responded to positively for use both while the course is running and for planning a revision of the course, and (v) a wide range of views was expressed about other details.

Author Keywords
Visualizations; Instructor support; e-learning; MOOCs; Massive open online courses.

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
In brick-and-mortar classrooms, instructors rely on face-to-face interaction with individual learners in lecture and office hours to understand how learners are doing in the course and how they interact with the course materials. Many recent Massive Open Online Courses (MOOCs) from providers such as edX and Coursera have enrolled tens of thousands of students per offering, with a few enrolling hundreds of thousands. At such scales, individual interaction with every student is infeasible and most interactions are through the software platform, rather than face-to-face. Fortunately, MOOCs' large scale and the fact they are offered via a heavily instrumented online environment, provide instructors with a rich source of information they previously lacked: instrumented activity from interactions with the e-learning platform.

Historically, data visualization has been an effective way to explore large datasets in which identifying interesting patterns is more productive than scrutinizing individual data points. Since MOOCs are relatively new, little work has been done on visualizing the rich sources of information available in them; current MOOC platforms offer only a small set of visualizations of basic quantitative information.

To help explore this space, we investigate it from two angles. First, we implemented a prototype instructor dashboard for the edX platform called the Metrics Tab (see Figure 1) that is currently available only to a small number of test users. Second, and the focus of this paper, we administered a survey to investigate the following questions:

1. What information sources do MOOC instructors prefer to help identify key trends and behaviors in both student performance and student interaction with course content?
2. Which of these sources are most useful to instructors during the three phases of: course preparation, course administration, and course postmortem?
3. How should these sources be presented so we may develop tools and visualizations instructors will find most useful?

The survey was answered by 92 MOOC instructors. Survey questions also include visualizations of information sources, two of which were modeled after those in the Metrics Tab, as well as three additional designs. Instructors were asked to judge the understandability and usefulness of each design.

The results support the following primary findings:

1. Quantitative data sources such as assignment grades are not enough: understanding discussion forum activity was of interest to 97% of those surveyed that answered questions on

L@S'14, March 4–5, 2014, Atlanta, Georgia, USA.
ACM 978-1-4503-2669-8/14/03.
http://dx.doi.org/10.1145/2556325.2566246

Information Visualized	Related Work
Performance: Grades on assignments, cumulative performance on problems for a particular concept	[1, 6, 15, 17, 18]
Access and Activity Patterns: What, how much, and when content has been opened, how long a student stays on a piece of content, student navigational path through the content, when a student turned in an assignment	[1, 2, 6, 8, 11, 12, 15, 17, 18, 20]
Forum Discussions: Author of a post, when the post was made, structure of follow-up posts, how many posts a student made, how many follow-up posts are in threads each student made, number of posts read by a student	[7, 8, 15, 17, 18]
Student Demographics: Location, reason for taking the course, age, learning style	[12, 20]

Table 1. Types of information and related work that visualizes each.

use of information sources. This is despite a lack of related work on visualizing discussion forum activity at scale and despite previous work showing that forum use is typically limited to a small percentage of students who are not necessarily representative of the overall enrollment [3, 4].

2. Instructors do not think chat logs are a valuable information source for understanding student behavior.

3. By and large, MOOC instructors want the same sources of information as instructors of smaller-scale distance learning courses, as evoked by earlier surveys.

4. Respondents reacted positively to mockups of both previously-used and novel visualization techniques, indicating they would use these to monitor a running course and to review materials when preparing for a new offering, but were less likely to use them in preparing new material.

5. Instructors expressed widely varying views on the types of data and visualizations they would find useful: some preferred data and visualizations that would support quantitative analysis such as correlation, others conducted courses focused more on discussion than quantifiable grades and therefore quantitative analysis is not useful, and so on.

Below we present related work, describe the survey procedure, describe the visualizations, present the results, discuss the ramifications of these results, and conclude with recommendations for future work for the design of monitoring interfaces for MOOC instructors.

RELATED WORK

Instructor surveys

Monitoring student learning has been promoted as a best practice in the education literature since the 1970s [5]. Two surveys of e-learning instructors, one in 2003 by Mazza et al. [16] (98 participants) and another in 2006 by Zinn et al. [21] (49 participants), agreed broadly on several points. Respondents stated that the most important phenomena to monitor are individual students' performance, per-student performance compared to the class as a whole, common misconceptions shared by many students (as manifested by common wrong answers to exercises, for example), and activity patterns such as what material students look at, how many times, for how long, and whether the material viewed is consistent with the course schedule. Mazza et al.'s respondents also said that forum behavior was a valuable way to gauge participation, but email or chat data was not.

Visualizations of student information

Visualizations have been used as a form of educational data mining [19]. However, very little related work in visualizing student information has focused on MOOCs, and modern MOOC platforms such as edX and Coursera provide limited instructor-facing visualizations. Table 1 shows information categories prior work has commonly visualized, ranging from standard graphs to innovative designs.

Standard graphs used by prior work [1, 2, 6, 8, 11, 12, 15, 17, 18, 20] include: scatter plots, bar indicators, bar and stacked bar charts, line graphs, Cumulative Distribution Function line graphs, pie charts, and heat maps. These graphs are used by prior work in one of two ways: (1) to provide a set of visualizations showing different kinds of information or (2) as a supporting graph in a complex visualization.

The prior work that provides visualizations for multiple categories of information [1, 6, 8, 12, 15, 17, 18] usually has the goal of giving instructors an overall picture of their course's e-learning experience. These visualization systems include: Goldberg et al.'s [8] early WebCT [9] visualizations, Hardy et al.'s [12] e-learning tracking visualizations, Mazza et al.'s Coursevis [15, 17] and GISMO [18], Gaudioso et al.'s [6] visualizations for dotLRN and PDinamet, and Khan Academy's Coach monitoring system [1]. It is important to note, while these systems provide an overview of the course, they often are intended for courses of only tens to the low hundreds of students and visualize each student individually (such as show each student as their own row in a heat map). Therefore a majority of these visualizations would not scale to the size of a MOOC unless judicious filtering is applied first.

Innovative visualizations used by prior work usually use known visualization techniques in an innovative way. These include: directional and non-directional node graphs, three dimensional graphs, timeline spiral graphs, icons, and line graphs. Node graphs are used by Hardless et al. [11] to show a timeline of student activity, calling it an activity line. Williams and Conlan's [20] use a node graph to show navigational path through content. Finally, Gibbs et al. [7] use a directional node graph, with node placement conveying time, to show how forum posts relate to each other.

Mazza et al. [15, 17] also visualize forums with a three dimensional scatter plot that the user could explore in. The timeline spiral graph by Aguilar et al. [2] shows student access and activity patterns. This graph used mainly bar graphs for both supporting information and spiraled around a center where each 360 degree spin was an easily understood unit of time (*e.g.,*24 hours, 1 week, etc.). Icons are used by Khan Academy's Coach tool [1] to highlight points in bar charts when students earned badges. Two prior works that use line graphs in innovative ways are Hardy et al.'s [12] line graph with shading to depict a student's path through the material and Williams and Conlan's [20] line graph as a connected sparse scatter plot depicting a student's learning style.

Four interaction techniques used in the prior work include: sorting, filtering, drill down, and clustering. Sorting was usually available in any visualization that provides a tabular view of information, such as Goldberg et al.'s [8] WebCT tabular student views and Khan Academy's Coach tool [1] that shows students individually. Aguilar et al.'s [2] timeline spiral allows users to filter by time, activity, course, and student. Hardy et al.'s [12] e-learning tracking visualizations allow filtering by time and any subset of students. It also incorporates an understanding of course hierarchy, which provides an ability to drill down through this hierarchy. Gaudioso et al.'s [6] visualizations for dotLRN and PDinamet included a clustering feature that automatically groups students based on access patterns. It provided a way to view aggregate information of the students in the groups and compare these aggregates to each other. Huang et al. [13] also uses clustering in their node graph to show syntax similarity between student code submissions.

Evaluation by prior work mainly involved interviews and focus groups [7, 11, 15, 17, 18, 20] and usually reported a mix of both positive responses to the system and a need for future improvements. Three prior works did not include an evaluation section [2, 8, 12]. Gaudiosoa et al.'s [6] work with dotLRN and PDinamet considered the drop out/success rate of the classes before and after the visualizations were provided to the instructors and found a marked improvement. However there is no discussion of whether the improvement with the dotLRN system is due to the visualizations or the revamped material that happened at the same time. They also conducted a questionnaire looking at student and teacher satisfaction, finding a majority of both groups were satisfied with the course and system. Mazza et al.'s [15, 17, 18] work on Coursevis and GISMO performed the most thorough evaluation, looking at the system's extent of required functionality, effectiveness, efficiency, and usefulness through a combination of an experimental study, interviews, and a focus group. Their results are positive across all their criteria.

SURVEY PROCEDURE

We used SurveyMonkey to administer a survey estimated to take about 30 minutes. We identified 539 potential participants by collecting instructor names from web page of courses offered on the three largest MOOC platforms: edX, Coursera, and Udacity.

The survey consisted of five parts:

1. **Background information** about the instructors.
2. **Specific details about one MOOC**. If an instructor taught multiple MOOCs we asked them to choose one and answer all following questions in terms of that MOOC.
3. **Course Monitoring Goals** and asking which information sources help achieve the desired understanding.
4. **Mockups** of five different visualizations of information that may be useful for monitoring a MOOC and questions about their efficacy.
5. **Open-ended response** for additional thoughts.

The next section describes the mockups in more detail.

THE METRICS TAB AND VISUALIZATION MOCKUPS

The instructors were asked to evaluate the potential usefulness and understandability of five visualizations of source information for monitoring MOOC activity. Two of these visualizations were derived from designs in the prototype Metrics Tab, a new tab in the edX Instructor Dashboard. Figure 1 shows this visualization in detail.

Both the mockup and the implemented prototype visualizations were based on ideas from previous work and informal conversations and brainstorms with instructors before the survey was administered. The mockups we decided to use also served as a preliminary evaluation of the Metrics Tab.

Metrics Tab

The goal of the Metrics Tab is to provide instructors a quick to consume dashboard display of available information in their course. The Metrics Tab separates the course's information by section and shows the same dashboard display seen in Figure 1 for each section. The section was chosen as the level of granularity because edX usually uses a section to contain a week's worth of material, with subsections allowing further division of the week's content.

The left grey bar chart shows how many students opened each subsection in the section; that is, viewed at least some of the content in that subsection at least once. When the user hovers the cursor over a bar in this graph, the name of the subsection and exact number of students that opened that subsection will appear in a tooltip, as seen in Figure 1.

The upper right red and green stacked bar graph shows the grade distribution for each problem in the section. It shows every problem regardless if the problem is included in the students' course grade or not. If students are allowed to submit an answer multiple times, as is common in MOOCs, it only shows the grade for their last submitted answer (since the last submitted answer is used when calculating a student's grade). For a given bar in the graph the color represents the grade for all the students in the bar, and the height is how many students received that grade. The color gradient for grades, seen to the left of the graph, goes from red, grey, to green[1], mapping to 0, 50, and 100 percent respectively. Hovering the cursor reveals the instructor-defined description of the problem, the number of students in the bar, their percentage grade, number of points earned, and number of possible points.

Finally, the bottom right blue stacked bar graph shows the distribution of number of attempts per problem in the section.

[1]This color scale is inappropriate for red-green color-blind viewers, and so in future iterations will be changed.

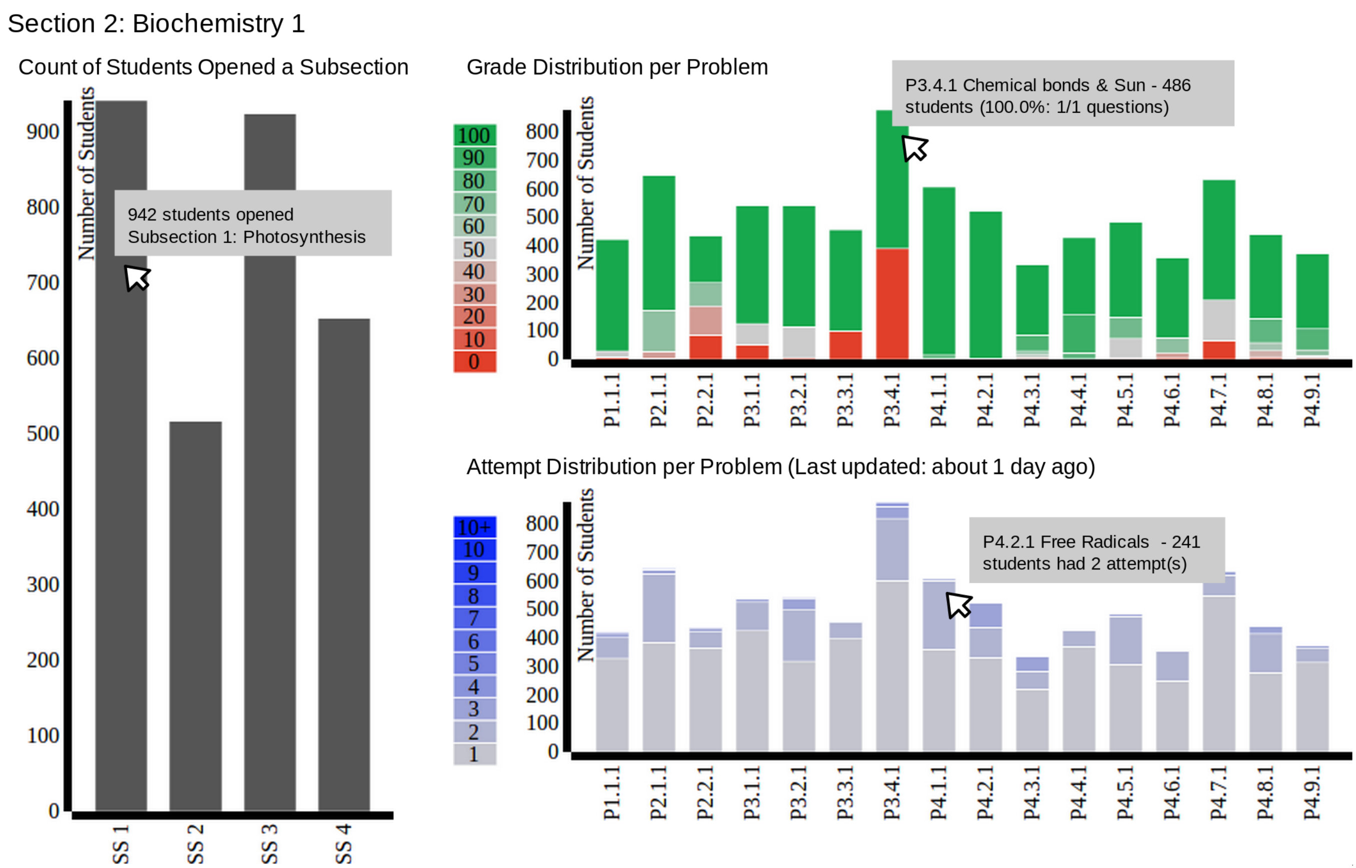


Figure 1. Mockup of a single section in the prototype edX Metrics Tab with a tooltip visible for each graph.

On both edX and Coursera, MOOC instructors can choose the number of times students may attempt each problem. The color gradient is grey to blue, mapping from 1 attempt to 10+ attempts. Students that attempted more than 10 times are grouped together because some problems allow unlimited attempts, which students do take advantage of. Hovering over a bar reveals the instructor-defined description of the problem, the number of students in the bar, and the number of attempts.

Visualization Mockups

The five mockups that were shown to instructors in the survey are shown in Figure 2. The callout bubble in each mockup represents what will be seen if the user hovers the cursor over that or a similar part of the graph. Mockup 2(a) is a boxplot diagram of grade distributions, included because it is a standard visualization. The tooltip shows the name of the homework assignment or other assessment item, the high and low scores, the median score and the 75th and 25th percentile scores. Mockup 2(b) is very similar to the upper right graph in the Metrics Tab and Mockup 2(d) is similar to the lower right graph of the Metrics Tab. Mockup 2(e) is similar to Mockup 2(d) but shows views of materials rather than attempts at homework problems. Finally, Mockup 2(c) shows two line graphs of forum usage data: number of new posts per day and number of views per day. The tooltip is for both graphs. On hover the points with the same date are highlighted. The tooltips text includes the date, the number of posts for that day, the number of posts viewed for that day and the titles of the most popular posts.

In the survey, each mockup in Figure 2 included a description on how to read the graph and any interactions with it.

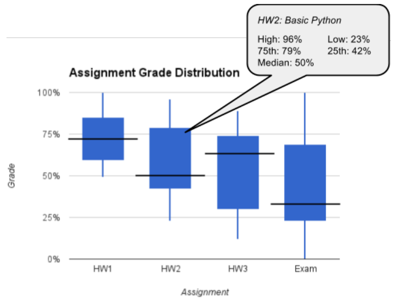
Participants were asked to provide Likert responses to (a) whether the mockup is useful and (b) whether it is easy to understand. Next we asked when the instructor might use it: (1) when preparing new material, (2) when preparing by reviewing past courses, and (3) while the course is running. We also ask if they have any other comments about the mockup (open-ended response).

SURVEY RESULTS

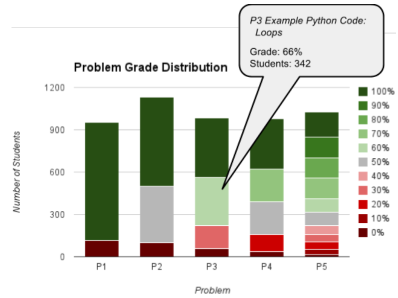
Of the 539 instructors solicited, 92 instructors (17%) started the survey and 67 (73%) completed it. Of the 91 instructors that chose to answer the question on gender, 73% identified as male, 25% as female, and 2% chose not to specify.

Characteristics of Courses

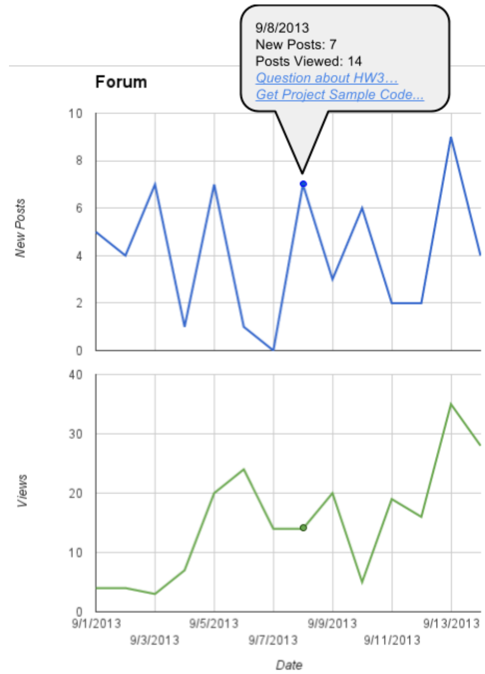
Of those instructors who ran a MOOC, more than two thirds had done so only one time, while 13% had done so twice (see Table 2). That said, many of these instructors are experienced at large in-person courses; 31% said they had run courses with greater than 250 people more than 4 times, and another 25% had done so 3 or fewer times. 85% of survey respondents reported creating one MOOC, 9% created two,



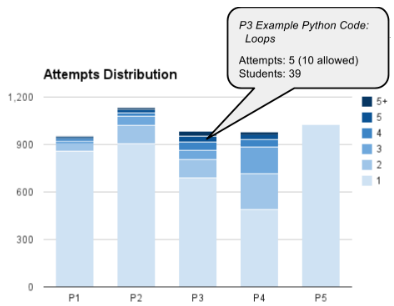
(a) Boxplot Assignment Grade Distribution



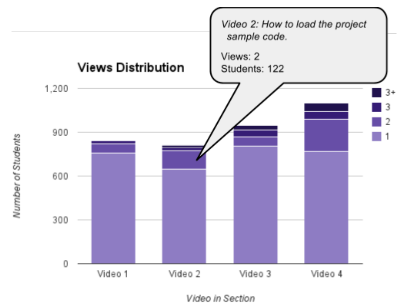
(b) Stacked Bar Graph Grade Distribution



(c) Line Graph Forum



(d) Stacked Bar Graph Attempts Distribution



(e) Stacked Bar Graph Views Distribution

Figure 2. Mockups shown to survey participants. The call out bubbles represents what will be seen if the user hovers the mouse over a part of the graph.

#	Run (N=92)	Created (N=91)
1	67%	85%
2	13%	9%
3	4%	1%
4+	9%	1%
N/A	6%	4%

Table 2. Survey respondents' MOOC experience for running and creating 1 to 4+ MOOCs.

	100s	1,000s	10,000s	Total
Engineering	3	10	7	20
Science	2	25	10	37
Humanities	1	8	10	19
Other	1	3	7	11
Total	7	46	34	87

Table 3. Cross between the estimated number of students in the course and the course's area.

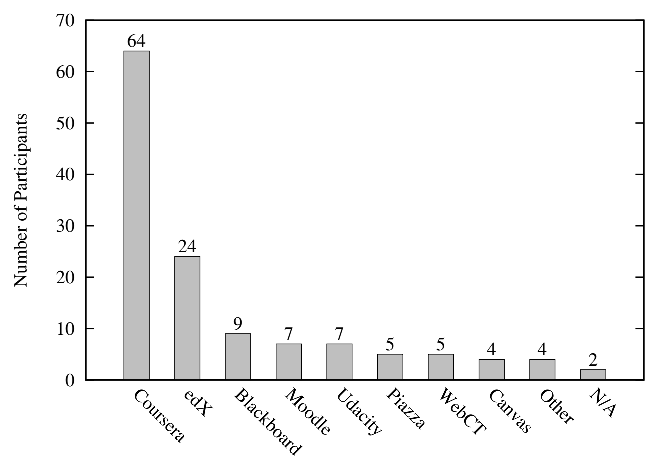


Figure 3. Platform usage statistics.

one individual created three, and one created four courses (see Table 2). Instructors have used a wide range of platforms, with Coursera and edX being the most frequent; Figure 3 shows the usage counts of the others reported. "Other" refers to platforms the instructors provided to us in the survey, which include Google, Desire2Learn, and an institution-specific MOOC platform.

For the questions that followed, if instructors had taught more than one MOOC, they were asked to choose one and answer in reference only to it. 91% of the courses completed with 1,000s to 10,000s of students. Table 3 shows the estimated number of students in the course at time of completion crossed with the subject matter of the course. The courses marked "other" include social sciences, business, education, and interdisciplinary studies. Interestingly, humanities courses were frequently among the largest MOOCS.

Course Monitoring Goals

We asked the survey participants to consider nine different tasks or goals they might have when running or planning a MOOC, summarized in Table 4. We asked instructors to assess ten information sources in terms of their efficacy for these nine goals, rating them in terms of if they currently *use*,

1	Problems with the current assignment.
2	Struggling students and what they are struggling with.
3	The difficulty of an exam problem.
4	Appropriateness of course difficulty level for students.
5	Most engaging content for the students.
6	Most difficult parts of the course.
7	Effectiveness of teaching assistants.
8	How to improve presentation of a topic.
9	Content students considered least interesting.

Table 4. Short descriptions of Course Monitoring Goals.

F	Discussion Forum
CS	Class Survey
SD	Discussion with Students
TA	Ask the teaching assistants
AG	Assignment grades
PAn	Student answers to problems
VP	Student's view pattern of online content
PAt	Number of times students attempt a problem
SCQ	Grades for self-check questions
CL	Chat room logs

Table 5. Resources Potentially Used for Course Monitoring.

would use if available, or *do not use/would not use* that resource. Asking about usage is different than prior work [16, 21], which asked survey participants to rate the level of interest or importance of an information source. We also purposely chose resources that instructors are likely familiar with to reduce the need for the instructors to guess if a resource is useful. The resources asked about are shown in Table 5.

The main findings from this section of the survey are:

Qualitative information is important. Figure 4 shows the raw counts of responses across monitoring goals for each information source. (Visual inspection did not reveal significant differences when subsets of courses were examined by area, but we did not do statistical tests to confirm this.) The lower segment (blue, solid border) indicates those who currently use this resource or would use it if available. The top segment (red, dotted border) indicates the counts of those who would not or do not use this resource over all tasks. For some questions, participants chose to answer usage for a subset of information sources. We assign the center (green, dashed-lines border) bar for the not-answer responses in order to show the relative rate of response.[2] The figure is ordered by highest to lowest raw response counts of use or would use if available.

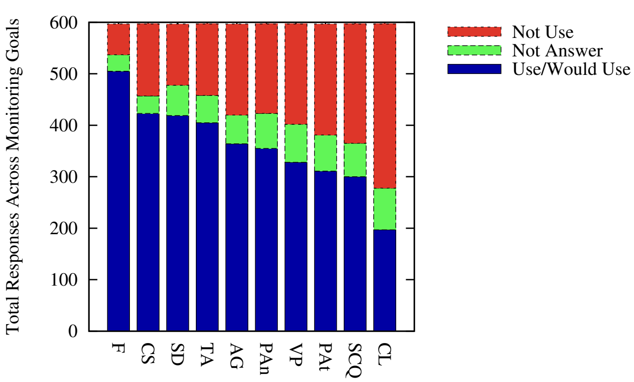


Figure 4. For each information source, the number of participants who use it or would use if available (combined into a single category), do not use, or did not answer. Did not answer means that the participant chose an answer for a subset of the information sources for that Monitoring Goal. It is shown here to see the relative rate of responses.

Figure 4 shows that across all tasks, discussion forums are seen as most often useful and chat logs are seen as least useful. In particular, the preference of forums over class surveys (the second-highest-used information source) is significant (Fisher's exact test, $p < 0.001$) and so is the difference between chat logs and self check questions (the second-lowest-used information source; Fisher's exact test, $p < 0.001$).

If we visualize the responses by Course Monitoring Goal, the pattern of response suggests a grouping as shown in Figure 5. The figure suggests groups of Course Monitoring Goals within which the relative usefulness of different information sources is more uniform than it is outside the group boundary. To make comparison easier, each separate stacked bar graph is the same as Figure 4, except it is the percent of instructors that answered for that Course Monitoring Goal instead of the raw counts. Results are not significantly biased because each question was answered by 59 to 75 instructors.

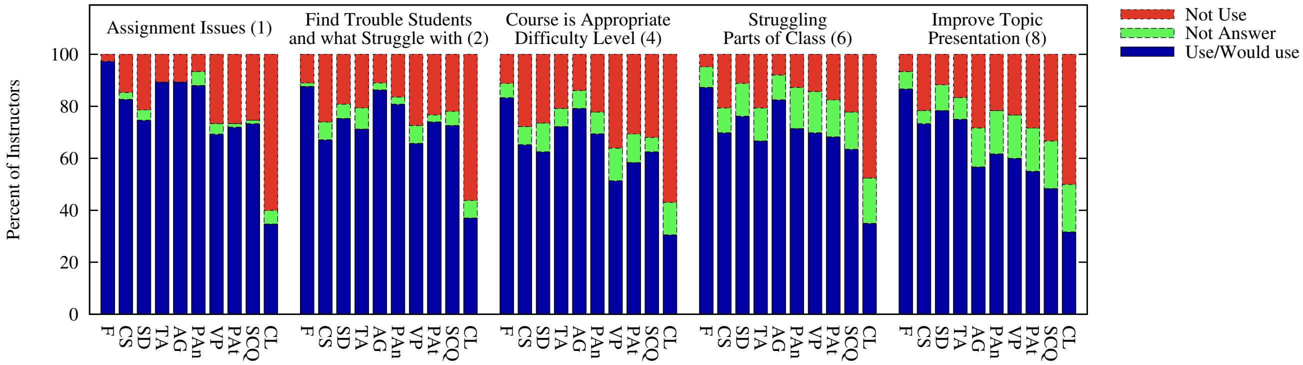
The first group of five Monitoring Goals (1, 2, 4, 6, and 8, in Figure 5(a); the numbers correspond to Table 4), could be characterized as quantitative questions about course material difficulty or presentation of course materials. For these Monitoring Goals all information sources but chat logs have 97% to 48% of question respondents say they use/would use the information source. Chat logs have only 37% to 31% question respondents say they use/would use it.

The second group of three Monitoring Goals (5, 7, and 9, in Figure 5(b)) could be characterized as qualitative assessment of student engagement or instructor effectiveness. In these goals participants were most enthusiastic about the "softer" information sources such as forums, discussions with students, class survey, and discussions with TAs. The percent of instructors saying they use/would use these information sources range from 94% to 46%. While there was much less enthusiasm for quantitative performance such as assignment grades, problem answers and attempts, and self check question grades, use/would use range from 41% to 15%. View pattern usage is the least similar between the goals, where it is used very little for the TA effectiveness goal, but used much more for the other two engagement goals.

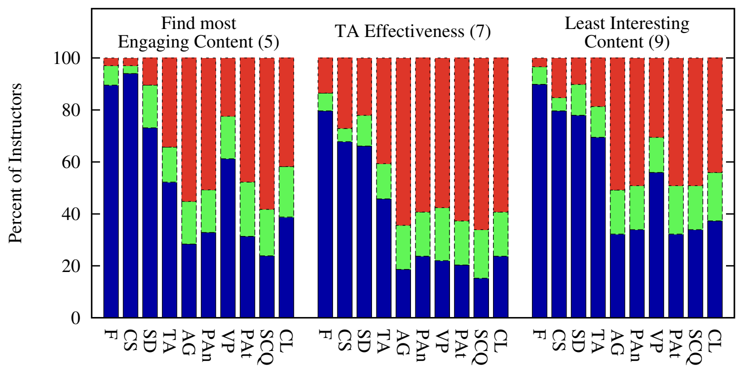
Monitoring Goal 3 – gauging exam problem difficulty in Figure 5(c) – does not display a similar pattern to the others, with no clear winners among the information sources.

Chat room logs are rarely used and considered unimportant. Chat room logs are more not used than used. This can be seen from the earlier discussion of Figure 4 and looking at

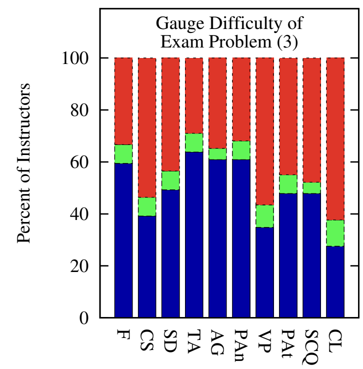
[2] Most participants completed the entire survey, 73%. However, because this section contained 9 questions requiring answers for 10 resources, some participants became fatigued (as indicated by their free-text comments) and either skipped portions of this part of the survey or did not complete the survey beyond this point.



(a) Information Source Usage Distribution Group 1



(b) Information Source Usage Distribution Group 2



(c) Information Source Usage Distribution Group 3

Figure 5. Percent of participants that answered for each usage option, as well as the percent that answered part of the question but not for that information source option. The Monitoring Goals are grouped based on their usage distributions. Each goal has a short description and the number corresponds to Table 4. Letters along the x-axis stand for the information source, see Table 5.

chat logs across all questions in Figure 5. Across all Monitoring Goals the percent of instructors saying they use/would use chat logs ranged between 39% and 24%. We exclude chat logs from the remainder of the discussion.

Instructors' opinions of what is useful largely confirm earlier surveys. Respondents' opinions of what information they would find useful is mostly consistent with the Mazza et al. [16] and Zinn et al. [21] surveys. In those surveys, the most important information concerns overall student performance relative to the class and information about what materials students interacted with and for how long (activity/view patterns); a majority of respondents also identified that information as useful. Respondents also agree with Mazza et al. that qualitative information from forum postings is important, but analysis of chat logs is not. However, respondents of both prior surveys placed higher importance on viewing per-student performance information and per-student mastery information. We speculate that such information is less useful in MOOCs, in which attention to individual students is rare.

Open-ended portions of the survey revealed a wide range of instructor views. Most survey questions, and each section of the survey, solicited open-text comments. The responses showed instructors' preferences ranging from simple numbers with no visualization to very complex data analysis. Complex analysis tool requests included A/B testing, correlational analysis, auto clustering of students by instructor-chosen parameters, and detailed view pattern information including paths through material.

An interesting dichotomy arises between those courses where quantitative assessment is foremost and more experientially-oriented courses. Several instructors stated they were not interested in grades, problem answers, and other performance based metrics because their goal was to provide students with a learning experience and not the ability to quantitatively prove they learned the material. These instructors stated they ran their course based on discussions and team interactions.

Some instructors were not worried about certain monitoring goals. One instructor stated there were too many students to worry about finding struggling individuals. Two instructors said course difficulty was not a concern. One stated course difficulty was fixed at the beginning and could not change while the course was running, while the other said they structured their course to work at multiple difficulty levels.

Responses to Mockups

Before going into detail of the mockups responses, it is important to note a caveat to these results. As a reminder, the survey asked participants if they considered each mockup useful and easy to understand and to predict when they would use the mockup. Since the survey asked participants what

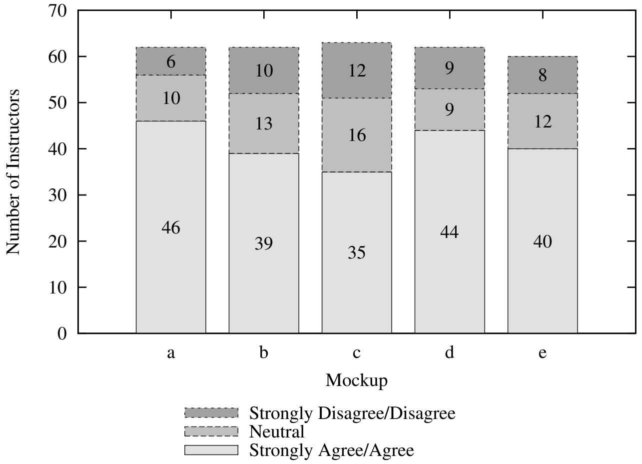


Figure 6. Likert scale responses to the statement "This visualization is useful." Letters correspond to Figure 2's subfigures.

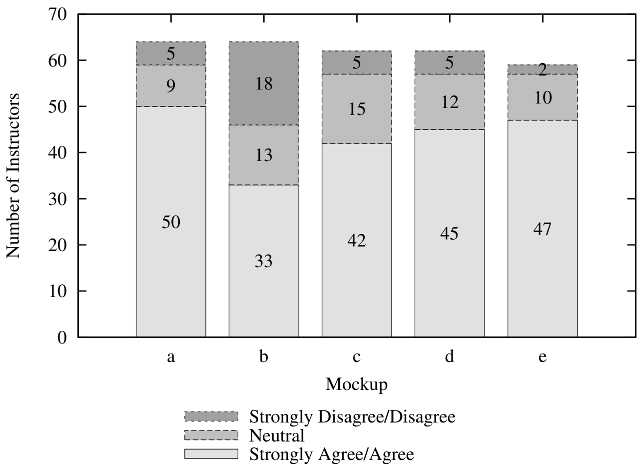


Figure 7. Likert scale responses to the statement "This visualization is easy to understand." Letters correspond to Figure 2's subfigures.

they think they will like and do, as opposed to what they actually liked and did, there are limitations to these results generalizability because what a person thinks they will like or do does not necessarily match what they will actually like or do.

The results of the mockup section are:

At least a majority of instructors considered each mockup useful and understandable. Figures 6 and 7 show participants' responses to the visualization mockups. The familiar box plot (Mockup 2(a)), when applied to student grades, was most often viewed as useful (74%) and understandable (78%), followed closely by visualization (Mockup 2(d)) showing number of student attempts at assignments (71% useful and 73% understandable). The number of times materials were viewed (Mockup 2(e)) was also considered useful information by two thirds of respondents (66%). A number of people expressed concern that the stacked bar visualization of the grade distribution (Mockup 2(b)) was difficult to understand, with only 52% agreeing that it was easy to understand. The visualization of the forum usage (Mockup 2(c)) was also

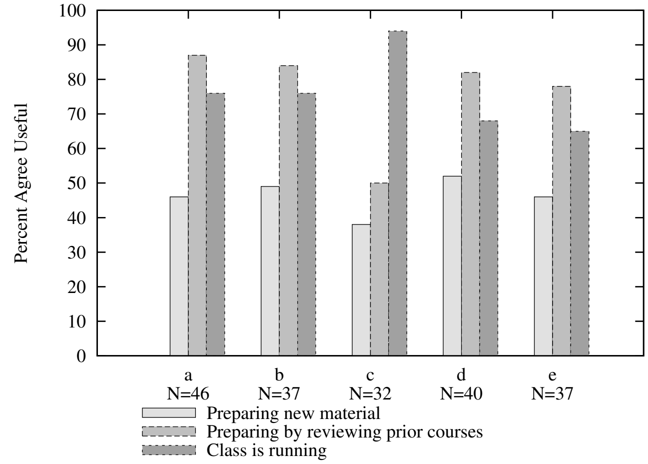


Figure 8. Responses to the question "When would you use this mockup," with choices in terms of three phases in a "MOOC cycle": preparing new material, preparing by reviewing previous course runs, and while the class is running. Letters correspond to Figure 2's subfigures.

not overwhelmingly supported, with only 55% of respondents agreeing or strongly agreeing that it was potentially useful.

There is a relative lack of interest in the forum visualization. Only 55% of instructors stated the forum usage visualization would be useful. Comments about this visualization stated it is not fine-grained enough, it is not more useful than existing statistics, the number of posts is not the useful indicator, and up-and-down votes are more important indicators.

Of those that considered the visualization useful, they would mainly use the visualizations while the course is running and all visualizations but forums after the course is over while preparing for a future offering. For those who did indicate that a given visualization was potentially useful, they were asked to indicate which circumstances it would be best used. Figure 8 presents the results; participants could mark more than one choice in each case. In all cases, at least two thirds of respondents who found the visualization useful wanted to use it while the course was running. And by an even larger margin, instructors wanted to see the visualization for every design except the forum visualization (c), when reviewing a past course in preparation for a future offering.

Relatively fewer participants considered the visualizations useful while preparing new course material. 38% to 52% of respondents indicated preparing new material for a course would be a good use of the visualizations.

METRICS TAB USAGE EXPERIENCES

A variation of the Metrics Tab was released to a small number of test users. This variation included the open subsection count graph (Figure 1 left, grey) and grades graph (top right, green and red); the attempts graph was not available. We interviewed two users that used the Metrics Tab while their course was running. Also, at the time of publication, we became aware of another publication that used the Metrics Tab [10]. We report the Metrics Tab usage experience below.

One user we interviewed was the TA of a MOOC that started with about 8,000 students and ended with about 500 completing the course. The Metrics Tab was one of the TA's primary methods for tracking student activity. The open subsection count graph was used to monitor how many students were still active in the course and if they are looking at all of the content. The grades graph was closely monitored to see how many students were doing the problems and which problems might have issues that needed to be resolved, such as input errors or ambiguities in the question,

The second interviewee was an instructor that ran a small online course of about 100 students. This instructor also used the open subsection count graph to see what content the students looked at. This instructor shared a story in which the Metrics Tab drew their attention to a student error; at the beginning of the course many students were unaware of any but the first subsection and had to be informed that there were other subsections. The instructor also liked the grades graph and used it to monitor the students' lecture quiz grades.

The final usage experience of the Metrics Tab is from Grover et al. [10]. They used the edX platform to teach an introductory computer science middle/high school course and reported on the pedagogy of their course and leveraging the Instructor Dashboard for curriculum assessment. They used the Metrics Tab to monitor quiz data, specifically using the grades graph to find what content the students found difficult and what content needed revision.

All three experiences confirm the survey finding that student performance information is important. The interviewees use of the open subsection graph aligns with the survey that student viewing patterns are important. Although only 52% of those surveyed found the Mockup 2(b) (which is based on the Metrics Tab's grades graph) easy to understand, neither of those interviewed had trouble interpreting the grades graph. The two interviewees were both engineers and so may be more familiar with reading graphs than other instructors.

DISCUSSION

The survey results show that discussion forums were the most frequently preferred source of information, across monitoring tasks, suggesting that effort should be invested in making forums more useful for students and more effective for providing information to instructors.

It is important to bear in mind, however, that prior work has found that only a small proportion of MOOC students are active on forums [3, 4], and therefore forum posters are not necessarily representative of all students in the course. Most likely instructors are aware of these limitations, and this may be why other methods for eliciting an understanding of students' views – student surveys, student discussion, and asking TA's for feedback – follow discussion forums as the perceived most useful information resources.

Notwithstanding these caveats, research to improve forums could significantly aid instructors' goals. For example, better methods to automatically group similar issues together, and to alert students to previously posted issues as they type, will help consolidate issues for the benefit of both instructor and student. User interface improvements can also help. Currently some forum tools, such as Piazza, allow the instructor or students to mark individual issues as resolved, but do not make it easy to group together a set of posts and mark them as similar and then resolved. A more "dashboard-oriented" view of forum posts, oriented towards the instructor, could be a significant time-savings improvement both for monitoring issues and topics in the forum and for processing posts as they are responded to by the instructor and teaching assistants.

This idea can be taken still further to create more of a "bug report" or "issues tracker" type interface approach to teaching a MOOC, similar to how problems are tracked with software engineering projects. As the instructor or teaching assistants learn about problems, via forum, survey, quantitative view such as low scores on a homework problem, the issue could be entered into this interface. Quick surveys or polls could be issued to see if the perceived gaps or problems are widespread and the results entered into the tool. After the correction is made, the problem could be marked as resolved.

Since surveys are private, those students who are not comfortable posting on the forum may be more willing to answer a survey and thus have their views expressed.

Finally, automated methods can be used to find which students appear to be struggling and send them survey questions or encourage them to read the posts on the forum or post their own questions, which will then be seen by the instructors.

A potential drawback of the work reported here is it primarily asked instructors about familiar information sources. Researchers are developing very sophisticated log analysis tools (*e.g.,* [13, 14]) that can produce profiles of student behavior that could be surfaced to the instructor in innovative ways. Future work must investigate the efficacy of these approaches.

Another drawback is MOOC instructors, coming primarily from in-person class backgrounds, may have preferences for technologies that work well in those environments and against those that do not, such as chat rooms. It may be the case that online chat will work better in MOOCs. More generally, after being exposed to new techniques in action, instructors may form different opinions.

The open-ended comments written by respondents revealed an interesting diversity of views that indicated what is useful to one instructor may not be as useful to another. Instructors ranged from being very pressed for time and wanting to see just a few summary numbers to wanting complex correlational analysis. Some courses are administered with a heavy quantitative evaluation focus, whereas others are more oriented around discussion. The emphasis on student grades in the survey seemed inappropriate to the latter instructors. This diverse range of instructor desires and course styles suggests that what is most useful and effective could be instructor- or course-specific and that the ideal MOOC data visualization should be flexible enough to meet these different needs.

CONCLUSIONS

This survey of 92 MOOC instructors confirmed the findings of prior surveys of instructors of conventional online courses, which found that instructors value seeing student performance, activity patterns such as what materials students look at, and forum behavior to gauge participation. A standard boxplot of distribution of course grades was seen as both understandable and useful by a large majority of respondents, as was a novel design in which stacked bar charts show number of repeated attempts at solving problems.

However, for those who wish to design visualization for MOOC instructors, a major takeaway from this work is views of quantitative measures are not sufficient. Rather, instructors believe they need to hear what students have to say, be it from discussion forums, student surveys, or the impressions of their teaching assistants, for the full range of course monitoring goals. Thus future work for instructor tools should focus on how to obtain the thoughts of a wide range of students taking the course, and how to summarize and present this information to the instructor in a useful manner.

ACKNOWLEDGMENTS

This work was supported by the UC Berkeley Chancellor's fellowship and the National Science Foundation Graduate Research Fellowship under Grant No. DGE 1106400.

REFERENCES

1. Khan Academy coach demo. `http://www.khanacademy.org/coach/demo`.
2. Aguilar, D. A. G., Therón, R., and Peñalvo, F. J. G. Semantic spiral timelines used as support for e-learning. *J. UCS 15*, 7 (2009), 1526–1545.
3. Breslow, L., Pritchard, D. E., DeBoer, J., Stump, G. S., Ho, A. D., and Seaton, D. Studying learning in the worldwide classroom: Research into edX's first MOOC. *Research & Practice in Assessment 8* (2013), 13–25.
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* (2014).
5. Cotton, K. *Monitoring student learning in the classroom*. Northwest Regional Educational Laboratory, 1988.
6. Gaudioso, E., Hernandez-del-Olmo, F., and Montero, M. Enhancing e-learning through teacher support: two experiences. *Education, IEEE Transactions on 52*, 1 (2009), 109–115.
7. Gibbs, W. J., Olexa, V., and Bernas, R. S. A visualization tool for managing and studying online communications. *Educational Technology & Society 9*, 3 (2006), 232–243.
8. Goldberg, M. W. Student participation and progress tracking for web-based courses using WebCT. In *Proceedings of the Second International NA WEB Conference* (1996), 5–8.
9. Goldberg, M. W., Salari, S., and Swoboda, P. World wide webcourse tool: An environment for building www-based courses. *Computer Networks and ISDN Systems 28*, 7 (1996), 1219–1231.
10. Grover, S., Pea, R., and Cooper, S. Promoting active learning & leveraging dashboards for curriculum assessment in an OpenEdX introductory CS course for middle school. In *Learning @ Scale, Work in Progress*, ACM (2014).
11. Hardless, C., and Nulden, U. Visualizing learning activities to support tutors. In *CHI'99 extended abstracts on Human factors in computing systems*, ACM (1999), 312–313.
12. Hardy, J., Antonioletti, M., and Bates, S. e-learner tracking: Tools for discovering learner behavior. In *The IASTED International Conference on Web-base Education* (2004).
13. Huang, J., Piech, C., Nguyen, A., and Guibas, L. Syntactic and functional variability of a million code submissions in a machine learning mooc. In *AIED 2013 Workshops Proceedings Volume* (2013), 25.
14. Kizilcec, R. F., Piech, C., and Schneider, E. Deconstructing disengagement: analyzing learner subpopulations in massive open online courses. In *Proceedings of the Third International Conference on Learning Analytics and Knowledge*, ACM (2013), 170–179.
15. Mazza, R., and Dimitrova, V. Coursevis: Externalising student information to facilitate instructors in distance learning. In *Proceedings of the International conference in Artificial Intelligence in Education*, Sydney, Australia (2003), 117–129.
16. Mazza, R., and Dimitrova, V. Informing the design of a course data visualisator: an empirical study. In *5th International Conference on New Educational Environments (ICNEE 2003)* (2003), 215–220.
17. Mazza, R., and Dimitrova, V. Coursevis: A graphical student monitoring tool for supporting instructors in web-based distance courses. *International Journal of Human-Computer Studies 65*, 2 (2007), 125–139.
18. Mazza, R., and Milani, C. Exploring usage analysis in learning systems: Gaining insights from visualisations. In *AIED05 workshop on Usage analysis in learning systems*, Citeseer (2005), 65–72.
19. Romero, C., and Ventura, S. Educational data mining: a review of the state of the art. *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on 40*, 6 (2010), 601–618.
20. Williams, F. P., and Conlan, O. Visualizing narrative structures and learning style information in personalized e-learning systems. In *Advanced Learning Technologies, 2007. ICALT 2007. Seventh IEEE International Conference on*, IEEE (2007), 872–876.
21. Zinn, C., and Scheuer, O. Getting to know your student in distance learning contexts. In *Innovative Approaches for Learning and Knowledge Sharing*. Springer, 2006, 437–451.