"It's all about conversation:" Challenges and Concerns of Faculty and Students in the Arts, Humanities, and the Social Sciences about Education at Scale

TIFFANY WENTING LI, University of Illinois at Urbana-Champaign, USA
KARRIE KARAHALIOS, University of Illinois at Urbana-Champaign, USA
HARI SUNDARAM, University of Illinois at Urbana-Champaign, USA

As colleges and universities continue their commitment to increasing access to higher education through offering education online and at scale, attention on teaching open-ended subjects online and at scale, mainly the arts, humanities, and the social sciences, remains limited. While existing work in scaling open-ended courses primarily focuses on the evaluation and feedback of open-ended assignments, there is a lack of understanding of how to effectively *teach* open-ended, university-level courses at scale. To better understand the needs of teaching large-scale, open-ended courses online effectively in a university setting, we conducted a mixed-methods study with university instructors and students, using surveys and interviews, and identified five critical pedagogical elements that distinguish the teaching and learning experiences in an open-ended course from that in a non-open-ended course. An overarching theme for the five elements was the need to support students' self-expression. We further uncovered open challenges and opportunities when incorporating the five critical pedagogical elements into large-scale, open-ended courses online in a university setting, and suggested six future research directions: (1) facilitate in-depth conversations, (2) create a studio-friendly environment, (3) adapt to open-ended assessment, (4) scale individual open-ended feedback, (5) establish trust for self-expression, and (6) personalize instruction and harness the benefits of student diversity.

CCS Concepts: • **Human-centered computing → Empirical studies in collaborative and social computing**; *Computer supported cooperative work*; • **Applied computing → Interactive learning environments**; **Distance learning**; **E-learning**; *Arts and humanities.*

Additional Key Words and Phrases: learning at scale; distance learning; online education; open-ended courses; arts; humanities; social sciences; MOOC

ACM Reference Format:
Tiffany Wenting Li, Karrie Karahalios, and Hari Sundaram. 2020. "It's all about conversation:" Challenges and Concerns of Faculty and Students in the Arts, Humanities, and the Social Sciences about Education at Scale. *Proc. ACM Hum.-Comput. Interact.* 4, CSCW3, Article 216 (December 2020), 37 pages. https://doi.org/10.1145/3432915

1 INTRODUCTION

In this paper, we ask: *how can we effectively teach open-ended university-level courses—mostly from the humanities, the arts and the social sciences—at scale?* We define "open-ended" (OE) courses as those which involve a significant amount of content that: (1) incorporates introspection and criticism, (2) raises questions with more than a small number of "correct" responses, and (3) is

Authors' addresses: Tiffany Wenting Li, University of Illinois at Urbana-Champaign, USA, wenting7@illinois.edu; Karrie Karahalios, University of Illinois at Urbana-Champaign, USA, kkarahal@illinois.edu; Hari Sundaram, University of Illinois at Urbana-Champaign, USA, hs1@illinois.edu.

2573-0142/2020/12-ART216 $15.00
https://doi.org/10.1145/3432915

amenable to subjective interpretation. We define non-open-ended (NOE) courses as those that do not satisfy all of the above criteria.

This question is salient for two reasons. First, as colleges and universities continue their commitment to increasing access to higher education [37], they differ in their strategies for offering education online and at scale. These strategies range from increasing enrollment in tuition-based online programs using in-house solutions [67, 32], to offering free or low-cost Massive Open Online Courses (MOOCs) via commercial platforms [62]. Open-ended courses account for a significant proportion of university-level courses. In 2019, traditionally open-ended subjects accounted for one-third of MOOC course offerings [66].

Second, surprisingly, despite the different characteristics between open-ended and non-open-ended courses, research [61, 25] indicates that instructors teach open-ended courses at scale through video lectures, quizzes, assignments, and peer assessment, in a manner similar to non-open-ended courses on mainstream MOOC platforms (e.g., Coursera, Udacity, EdX). Prior studies [67, 26] find that instructors in traditionally open-ended subjects had lower satisfaction and had more trouble when offering online courses. These two findings more directly motivate our research. We wish to better understand the challenges and concerns—of university faculty teaching open-ended classes face, and of students who take these classes—arising from university courses offered at scale.

Offering higher education online and at scale, despite its seemingly imminent arrival at universities, is, in fact, the latest iteration of a storied effort of offering education remotely and at scale. Looking back at the development history of distance education, we see a continuous effort to lower the barrier and enhance access to high-quality educational resources [81]. Beginning in the 1880s, print-based correspondence programs sent students learning materials and received assignments via mail [49]. In the 1920s, radio enabled broadcasting learning to anyone with a receiver (e.g., UK BBC radio programs [35]). The Wisconsin School of the Air (WSA) via radio, an exemplary distance learning program founded in 1936, additionally insisted that local schools or communities host a local exhibit of student artwork from its "Let's Draw" arts course, to foster a democratic learning community [5]. The WSA later set up a multi-channel distance learning system that included broadcast media, mail correspondence, and the telephone [17]. This distance learning system evolved into the Open University model [42], offering more interaction channels and spanning more countries. Starting in 1960, the first generalized computer-assisted instruction system, Plato, emerged at the University of Illinois. It offered interactive, graphical coursework to local university students, local schools, and other universities across networked computers world-wide [21]. This trend toward multiple-channel, interactive, distance learning continued at other institutions with the help of satellite teleconferencing, the personal computer, and the Internet. By 2019, MOOCs, fueled by falling hardware and communication costs, grew to include more than 110 million students [1], 900 universities, and 13,500 courses around the world [66]. Large scale, lower-cost, degree-granting online programs that are enrolling thousands of students per program per year, such as the Online MS in Computer Science at Georgia Tech and the iMBA program at the University of Illinois at Urbana-Champaign, are gaining popularity [32].

We see that in education at scale, over time, the costs (e.g., for materials, communication, computing) have fallen, there have been improvements in the level of interactivity, and the course materials have become available worldwide. At the same time, we have seen significant progress in computerized testing and auto-grading in STEM-related fields [87, 86, 9, 48, 50], decreasing attendant evaluation costs, thus making them more amenable to scale.

Feedback and evaluation in open-ended subjects remain expensive. To scale open-ended subjects, infrastructure that reduces feedback and evaluation costs is essential. Such an infrastructure is vital

[1] excluding China

because, traditionally, we teach open-ended courses in small, in-person settings that support highly individualized feedback. For example, in the art & design, the National Association of Schools of Art and Design guidelines recommend a studio class size of 20 students [58].

There is a significant interest in reducing the costs of feedback and evaluation in open-ended subjects. PeerStudio [47], Juxtapeer [11], CritiqueKit [28], and Paragon [41] are examples of tools developed to improve the quality of peer critique on open-ended assignments. Work on auto-grading [56, 10, 4] short-answer [56], essay writing [4], and open-ended questions has produced important contributions [30]. On the other hand, to the best of our knowledge, there is little work on how to effectively *teach* open-ended subjects at scale. The heavy focus on tools for scalable individual feedback and evaluation of open-ended assignments, an important concern given their cost, has led to a surprising gap—we do not understand well the effective *teaching* practices for open-ended, university-level courses at scale.

Present Work: In this paper, we report on our mixed-methods study, conducted during April–December 2019, of teaching practices for open-ended, university-level courses, comprising surveys of educators and students and in-person interviews of instructors. We base our findings on courses at a large public university for two reasons: (1) universities offer a wider variety of open-ended courses, and (2) we were able to obtain easier access to instructor and student samples. The goal of our study is two-fold. First, we wish to better support educators of open-ended courses seeking to teach their courses at scale and online in a university setting, and to support students in these massive, online open-ended courses. Second, we wish to help the development of infrastructure that supports the teaching practices of these educators and the learning needs of students who wish to take these courses.

To the best of our knowledge, no prior work contrasts the teaching and student learning experiences of open-ended and non-open-ended, university-level courses. Thus, our first two research questions are:

RQ 1.1 How are university-level open-ended courses taught differently from non-open-ended courses in traditional university settings?

RQ 1.2 Based on conclusions from RQ 1.1, what can we infer about the critical pedagogical elements unique to teaching open-ended courses effectively compared to non-open-ended courses in higher education, from an instructor perspective?

RQ 2 From a student perspective, what are student preferences for how instructors should teach open-ended courses?

To answer RQ 1.1 and RQ 1.2, we surveyed university instructors to compare open-ended and non-open-ended course teaching practices. Then, to answer RQ 2, we surveyed university students to understand their preferences.

Our third research question directly addresses the main goal of the paper mentioned above:

RQ 3 What are the challenges and opportunities in teaching university-level, open-ended courses effectively, online at scale in a university setting?

To answer RQ 3, we conducted in-depth interviews with university instructors about their perspective and experience with large open-ended courses in-person or online, ranging from a size of 30 students in a studio class to a 300-students MBA class.

We gathered data (surveys, interviews) and performed analysis on this data, after approval from our university's Institutional Review Board. We recruited instructors and students for the surveys and the interviews from our institution, the University of Illinois in Urbana-Champaign, a large, public research-oriented (R1) university. We used quantitative and qualitative analysis techniques to analyze the collected data. For the surveys, we used statistical tests (Student t-test on interval data; two-proportion z-test for nominal data; Mann-Whitney U Test to compare ordinal data) and

thematic analysis [31] (on free-form survey text). We identified a set of critical pedagogical elements for effectively teaching open-ended courses by synthesizing the findings from the instructor survey and the student survey. We then used this set of five pedagogical elements to organize the themes from the instructor interviews, which we obtained via a grounded-theory-like approach [51]. Thus, while research questions RQ 1 and RQ 2 are essential on their own, they contribute to our analysis of RQ 3.

This work, with an emphasis on university-level, open-ended classes, contributes to the extensive body of work devoted to improving learning at a distance and at scale in two ways—identifying key pedagogical elements and future research directions.

Key pedagogical elements: Through a careful analysis of surveys of instructors and students, we identified five key pedagogical elements that distinguish the teaching and learning experiences in an open-ended course from that in a non-open-ended course in a university setting. To the best of our knowledge, this is the first systematic analysis of the unique pedagogical requirements of open-ended courses compared to non-open-ended courses.
These elements include: (1) having *extensive* interaction via multi-channels; (2) using open-ended assignments that allow for self-expression; (3) providing customized individual feedback through various channels; (4) fostering an open and unbiased learning and evaluation environment for self-expression; and (5) enabling personalized instruction as a requirement. These elements were either not valued as *necessary*, or not considered as *applicable* to non-open-ended courses. To sum up, supporting students' self-expression in the learning process was critical in open-ended courses. Our findings are significant because they may help instructors to teach better and prod institutions to better resource the teaching of open-ended classes either on a large scale or when the course has an online component.

Identify future research directions: We uncovered challenges and opportunities when incorporating the five critical pedagogical elements for open-ended courses online and at scale in a university setting, by synthesizing findings from semi-structured interviews of instructors with findings from the analysis of instructor and student surveys. The identified elements are specific to open-ended courses and provide a nuanced understanding of the challenges. In contrast, the challenges and opportunities of online education at scale identified in prior work are general and not subject-specific [20, 38, 27]. By comparing the needs identified by instructors against the state of the art learning-at-scale environment, we suggest six future research directions. These include: (1) facilitating in-depth conversations; (2) creating a studio-friendly environment; (3) adapting to open-ended assessment; (4) scaling individual open-ended feedback; (5) establishing an inclusive environment for self-expression; (6) personalizing instruction and harnessing diverse student-generated content. This contribution is significant because our guidance, which focuses on open-ended classes and which is informed by the practice of open-ended instructors and the learning experiences of students, may steer the development of technological infrastructure to better support the teaching and the learning of open-ended courses at scale, and online, in a university setting.

We organize the rest of the paper as follows. In the next section, we discuss related work. Then in Section 3, we describe the methods, including participant inclusion criteria, an overview of the surveys and interviews, and the corresponding analysis techniques used to address our RQs. In Section 4 and Section 5, we analyze in-depth the results of the survey (addressing RQ 1, RQ 2) and the in-person interview (addressing RQ 3). In Section 6, we identify open challenges and opportunities not yet addressed in prior work. Finally, we identify the limitations and future directions of this work in Section 7 and present our conclusions in Section 8.

2 RELATED WORK

In this section, we first describe prior work on the pedagogy of open-ended subjects versus that of non-open-ended subjects. Then, we present existing studies that uncover the current challenges of teaching open-ended subjects online and at scale. Lastly, we examine work focused on developing tools to support teaching open-ended courses at scale.

2.1 Pedagogy of Open-ended and Non-open-ended Subjects

Prior work has examined critical aspects of teaching and learning open-ended and non-open-ended topics, but there is little prior work contrasting them. Martin [53] identified the key aspects of teaching open-ended topics to include: active learning, discussion on value-laden topics, and innovative assessments. Overton [60] established the following key components of teaching and learning non-open-ended topics: the lecture with opportunities for student participation, laboratory or practical classes, small group tutorial, and industrial work experience. Prior work has identified common learning and teaching strategies applicable to both kinds of courses: inquiry-based learning [6, 60], team-based learning [73, 54], just-in-time teaching [69, 18]. However, we lack an understanding of how strategies differ for teaching open-ended and non-open-ended topics effectively.

To the best of our knowledge, our study is the first to contrast the pedagogical foci of the open-ended and non-open-ended topics. Discerning these contrasts helps us to identify requisite technological affordances to better support each topic.

2.2 Challenges in Teaching Open-ended Courses Online at Scale

The question of how to teach open-ended topics effectively online and at scale, has received limited scrutiny. Swan et al. [72] used a categorical tool to establish differences in pedagogical approaches between four non-STEM and nine STEM MOOCs. One of the differences Swan et al. [72] pointed out was that learning activities in non-STEM MOOCs are divergent, while those in STEM MOOCs are convergent. While these differences are important, they did not examine the effectiveness of the pedagogical frameworks used in existing non-STEM MOOCs. In contrast, we first determine the pedagogical needs of a university-level open-ended course, and then examine if the existing platforms support those needs.

There is limited work in examining the challenges in teaching courses in the humanities at-scale effectively. Evans and McIntyre [25] uncovered a challenge that many humanities MOOCs did not serve students with varying degrees of preparation and underprivileged students well using textual analysis. Evans and McIntyre [25] found that while most MOOCs in the humanities included "no background required; everyone is welcome" in the course description, most MOOC instructors in the humanities included complex reading materials and assignments that often required high school or some college experience. Instructors also expected students to have high-speed internet and disposable income. While Evans and McIntyre [25] had a specific challenge in mind to verify, our study has an exploratory nature and aims to uncover a broader range of challenges grounded in empirical evidence.

Prior work has examined the unique challenges in teaching the arts at scale. Peng [61] proposed that the unique needs of arts education include: active learning, personalized feedback, and comprehensive assessment. Peng [61] argued—using a case study of Kadenze, a pioneering arts-focused MOOC platform—that current MOOC platforms do not serve the needs of arts education. Peng [61]'s identified needs for the arts are reasonable, and in this work, we seek to ascertain critical elements common to teaching open-ended subjects, including the arts, at scale, through a careful examination of current teaching practice. In studies related to technological platforms for the arts, Xiong et al. [82] proposed a vision for an art education online platform leveraging big data,

incorporating personalized tutoring and virtual reality. Xiong et al. [82] envisioned a platform that complemented teaching of in-person art courses, and lacked a focus on scale.

2.3 Support and Solutions for Open-ended Courses Online at Scale

Recent work on supporting open-ended courses online at scale focuses on a critical aspect: evaluation at scale. PeerStudio [47], Juxtapeer [11], CritiqueKit [28], and Paragon [41] are tools that designed to improve the quality of peer assessment on open-ended assignments. Kulkarni et al. [47] proposed a peer review system, PeerStudio, that could provide peer feedback per request in 20 minutes on average. Fraser et al. [28] developed a mixed-initiative machine learning system, CritiqueKit, that can recommend prior feedback to instructors for them to reuse in assignments with similar mistakes. Cambre et al. [11] developed Juxtapeer, a peer review system that utilized peer work comparison in the feedback writing process and produced better quality peer feedback. Lastly, Kang et al. [41] created Paragon, a peer feedback system, that enabled feedback providers to incorporate example work to make their reviews more specific, actionable, and novel.

Prior work demonstrates the utility of data analysis techniques for auto-grading open-ended assignments. Mohler et al. [56] utilized semantic similarity measures to grade short-answer questions, and Brooks et al. [10] clustered the answers to assist instructors. Geigle et al. [30] provided a supervised-learning solution to grade and rank medical case assessment automatically, an example of complex, open-ended assignments. Balfour [4]'s work on automated essay scoring may find use on grading some open-ended assignments.

Talkabout, by Kulkarni et al. [46], focuses on another critical aspect of learning at scale—facilitating peer interaction. Talkabout utilized guided video-conferencing to encourage conversations among geographically diverse students. The work by Kulkarni et al. [46] is an exception—there is limited prior work on supporting peer interaction on open-ended courses at scale.

To summarize, in our survey of related work on open-ended courses taught at scale, supporting evaluation at scale emerged as the area of emphasis. However, tools that support the evaluation of open-ended courses implicitly assume that education at scale platforms have the technological affordances that support the *teaching* of open-ended classes effectively. We suspect that many requisite technological affordances do not exist since prior work [61, 25] indicates that instructors teach open-ended and non-open-ended courses in a similar manner on the mainstream MOOC platforms.

Thus, to better understand how to better support open-ended courses online and at scale, we need to contrast the teaching and learning of open-ended courses with non-open-ended courses. We develop this understanding through instructor and student surveys (RQ 1, RQ 2, Section 4). Then, we explore the challenges and opportunities instructors would face in teaching university-level open-ended courses online and at scale in a university setting via instructor interviews (RQ 3, Section 5).

3 METHODS

Now, we present an overview of our mixed-methods study that involved university instructors and students. We discuss the instructor survey in Section 3.1, followed by a discussion of the student survey in Section 3.2 and conclude with the instructor interviews in Section 3.3. Table 1 shows a summary of the research instruments and analysis approach.

3.1 Instructor Survey

We designed an online survey for university instructors to solicit common teaching practices in open-ended and non-open-ended courses in traditional university settings, both in-person and online. We used the survey to understand what instructors have been doing and what they

Table 1. Instrument and analysis methods used for each research question

RQ	Instrument	Analysis Method
RQ 1	Instructor Survey	Statistical tests (Student t-test, Two-proportion z-test, Mann-Whitney U Test); Thematic analysis
RQ 2	Student Survey	Statistical tests (Student t-test, Two-proportion z-test, Mann-Whitney U Test); Thematic analysis
RQ 3	Instructor Interview	Grounded-theory-like approach (Thematic coding; Pattern coding); Framework analysis

considered essential in teaching open-ended courses effectively. We administered the survey over four weeks in November–December 2019. We developed the survey to complement the instructor interviews, enabling us to cover a larger instructor population.

Participants: Qualified participants were instructors in universities and who had taught at least one course in the past three years. We reached out to 2400 instructors and received 121 responses from our home institution, a public research university. 37.2% of the participants identified themselves as women, and 62.0% as men. The sample contained 44.63% professors, 20.66% associate professors, 21.49% assistant professors, 12.40% lecturers or instructors, and 5.79% other. If the participant had taught at least one online course, the survey instructed the participant to answer the survey based on their most recent online course; if they had not taught an online course, their responses should apply to a recently in-person class.

Table 2. Distribution of course type in the sample (n = 121). The instructors categorized their courses as open-ended, online or of a large size based on a consistent rubric. Our sample has a balanced number of open-ended and non-open-ended courses.

Category	Yes	No
Open Ended	61	60
Online	40	81
Large Size	46	75[a]

[a] There were 30 medium sized courses and 45 small sized courses.

The 121 courses discussed in the instructor survey were from 64 departments. Examples of departments with open-ended courses included political science, philosophy, theatre, music, social work, and fine & applied arts. For non-open-ended courses, examples of departments included mathematics, advertising, geology, economics, aerospace engineering. Some departments included both open-ended and non-open-ended courses, such as philosophy, sociology, information sciences. From Table 2, we can see that the sample contains an even number of open-ended and non-open-ended courses. Our sample contains a smaller number of online courses compared to in-person courses. The sample is uniform across various class sizes.

Survey Overview: The survey comprised five sections covering demographic, basic course information, class mechanics, evaluation and their views on pedagogical elements. First, we asked for basic information about the participant, e.g., the affiliated department(s), academic rank, and gender. In the second section, instructors self-reported whether they had offered an open-ended course before. If yes, we asked them to focus on their most recent open-ended course in the survey. Otherwise, they chose their most recent non-open-ended course. The rest of the second section addressed general information of the chosen class, such as the learning goals, class size, and difficulty level. In the third section, we asked how instructors conducted specific components

of the course (when applicable) and how they incorporated instructor-student interaction and peer interaction in each of the components. We asked about three class components in this section: a lecture-based session, a discussion or seminar session, and a lab or studio session (detailed definition in Appendix A.1). The fourth section of the survey contained questions related to assignments, evaluations, and feedback for students in the course. Lastly, we asked for instructors' opinions on the importance of various pedagogical elements for achieving their learning goals.

Analysis: We compared the pedagogical choices made by instructors between open-ended and non-open-ended courses. We tried our best to match the samples of open-ended and non-open-ended courses along the dimensions that may influence pedagogy decisions between the two comparison groups. We checked the percentage of online courses, the distribution of different class sizes, the instructor's academic rank, gender, class difficulty level, and the highest learning goal categorized by Bloom's taxonomy [7]. The open-ended and non-open-ended course samples are very similar in all of these factors (please refer to Table 6 in the A.2 for details). The similarity suggests that the differences we found between the two groups may be attributed to the difference in pedagogy used in open-ended and non-open-ended courses.

We used quantitative and qualitative analysis techniques in our comparison. We used statistical tests (see Table 1) to compare the distribution of responses between open-ended courses and non-open-ended courses, for multiple-choice questions or fill-in-the-blank questions with numeric answers. We applied the familiar Student t-test on interval data. We used the two-proportion z-test for nominal data since all cells had more than five samples. We applied the Mann-Whitney U Test to compare ordinal data. To analyze free-form text questions, one of the research team members used thematic analysis [31] to convert them into categorical responses. For questions related to an existing analysis framework (e.g., categorizing learning goals by Bloom's taxonomy [7]), we used a deductive coding approach. For questions that did not relate to an existing analytical framework, we applied a reflexive inductive coding approach to extract the themes.

3.2 Student Survey

Between June and November 2019, we distributed an online survey to college students who had taken at least one open-ended university course in the past four years. We aimed to understand, from a student's perspective, the characteristics of the open-ended courses they enjoyed or disliked, to complement the findings from our instructor survey.

Participants: We conducted two surveys and received 55 responses from students who had taken at least one open-ended course. We distributed the first batch of surveys to students in the open-ended courses taught by instructors we interviewed. It resulted in nine participants. Due to the low number of responses, we reached out to 2000 students via a student sampling service by our home institution in the second round of recruitment. We recruited 46 participants in this second round. Among the 55 participants, 76% of them were undergraduate students, and the rest were graduate students. Among the undergraduate students, the percentages of Freshman, Sophomore, Junior, and Senior were 1.82%, 18.2%, 20%, and 36.36%, respectively. Nearly 85% of them had taken more than one university-level open-ended course. Approximately 40% of them had taken four or more open-ended courses. Among the 46 students for whom we have gender data, 70% identified themselves as women.

Survey Overview: We designed the student survey, beyond soliciting demographic and course information, to focus on interaction with the instructor and with peers. We began the survey by asking students to self-report the number of open-ended courses they had taken based on our definition of an open-ended course. If the student had taken only one open-ended course, we asked what they liked and disliked about the course and instructed them to answer the rest of the survey based on their experience with this course. Otherwise, we asked the student to choose a favorite

open-ended class, explain why the course was their favorite, and answer the remaining survey with this class in mind, unless otherwise instructed. The rest of the survey consisted of four parts. The survey started with demographic questions and basic information about the course. Next, we focused on the usage of instructor-student interaction and peer interaction in the course. In the third section, we asked about assignments, evaluations, and feedback for students. Finally, for students that had taken more than one open-ended course, we asked them to pick an open-ended course that they disliked and to identify reasons why they did not enjoy the class.

Analysis: We used approaches similar to those used in the instructor survey. We analyzed the distribution of multiple-choice responses and performed statistical analysis where applicable, as shown in Table 1. We summarized results from free-form text questions via thematic analysis using a reflexive approach.

3.3 Instructor Interview

To identify the challenges and opportunities of teaching open-ended courses online and at scale effectively, we conducted in-depth, semi-structured interviews with instructors that had taught more than one open-ended courses from our home institution.

Participants: To recruit participants, we shortlisted targeted individuals from 16 departments[2] based on the syllabus of courses they taught and reached out to 48 instructors. We used a theoretical sampling strategy [51], where we focused on recruiting instructors that satisfied specific criteria. In our case, we mainly recruited instructors that had taught large or online open-ended courses. We recruited a small number of participants by word of mouth. Between April to August 2019, we recruited 11 of them from nine departments. Among the 11 instructors, five were from different departments within art & design, two were from the philosophy department, and the other four were from architecture, computer science, advertising, and business administration. Courses in these departments covered the three major types of open-ended deliverables defined in the instructor survey. The art & design, architecture, and advertising courses focused on artifact creation. Philosophy courses focused on argumentation. We included the computer science and business administration instructors as their courses focused on finding solutions to open-ended problems.

Five participants identified themselves as women, and the rest identified themselves as men. Among the eleven were one professor, one associate professor, five assistant professors, and two lecturers. Two respondents were graduate students who taught courses independently. Three of them have taught fewer than five open-ended courses. Another six had taught five to 19 open-ended courses. Two others had taught 20 or more open-ended courses.

Interview Overview: We conducted the semi-structured interviews in-person; they ranged from 50 minutes to two hours. We identified four types of courses organized in a 2×2 grid based on class size (small vs. large) and interaction type (in-person vs. online). The instructor categorized the size of their class in relation to their home department average. All of the interviewees had experience teaching a small, in-person course. If the instructor had experience teaching the other three types of courses (small & online, large & in-person, large & online), we asked them to focus on the other three types in the interview and compare them with their experience teaching a small in-person course. We guided them to share the challenges and opportunities they encountered when teaching a large or an online course. Otherwise, instructors described challenges in a small in-person course and shared what they imagined the challenges and opportunities that may arise if

[2]Department of Graphic Design, Studio Art, Art Education, Fashion Design, Computer Science, Philosophy, Music, Architecture, Advertising, African American Studies, Sociology, Business Administration, Psychology, Special Education, Information Science, and English

they were to teach the same course at a larger scale or online. We discussed a total of 16 courses, among which five were small in-person courses, five were large in-person courses, five were large online courses, and one was a small online course. For instructors in art & design and architecture, the size of large classes in discussion ranged from 30 to 100 students. The sizes of large classes in other instructors' experiences were between 100 to 300 students.

The interviews focused on five perspectives, class components and content delivery, peer interaction, instructor-student interaction, assignment, and evaluation. We identified the five perspectives based on the critical aspects of a course introduced in prior literature [20, 57]. First, per Daradoumis et al. [20], we identified two overarching aspects of a course, delivery, and assessment. Then, consistent with the categorization by Moore [57], we further split delivery into the three types of interaction (student-content, student-instructor, student-student), and assessment into assignment and evaluation.

Analysis: For our analysis, we first employed a grounded theory-like approach [51], where we inductively generated themes that summarized the topics discussed in the interviews with a reflexive and iterative process. One of the research team members started with thematic coding [31] to compile a list of codes inductively, followed by a semantic approach to iteratively refine the themes of challenges and opportunities. We then grouped the themes into two groups based on the underlying cause of the challenge or opportunity, 1) being massive and 2) being computer-mediated. After using a grounded theory-like approach to generate the themes, we applied a framework analysis [29] approach to summarize them. As mentioned in the last subsection, we created a set of critical and unique pedagogy elements for effectively teaching open-ended courses by synthesizing the findings from the instructor survey and the student survey. We used this set of five pedagogical elements to organize the themes from the instructor interviews. We dropped the interview themes unrelated to the five elements.

4 SURVEY RESULTS—CRITICAL PEDAGOGICAL ELEMENTS FOR OPEN-ENDED COURSES

In this section, we describe how we arrived at the set of critical pedagogy elements for open-ended courses from the instructor survey results and the student survey results. In the first subsection, we present the results from the instructor survey. In the second subsection, we describe the results of the student survey. The last subsection summarizes and combines the findings from RQ 1 and RQ 2 to arrive at a comprehensive set of crucial pedagogical elements that differentiated the teaching and learning experiences in open-ended courses from that in non-open-ended courses. We conclude by identifying five elements that make for a positive experience in open-ended courses.

4.1 The Instructor Perspective (RQ 1)

In this section, we present our findings for our first set of research questions:

RQ 1.1 How are university-level open-ended courses taught differently from non-open-ended courses in traditional university settings?

RQ 1.2 Based on conclusions from RQ 1.1, what can we infer about the critical pedagogical elements unique to teaching open-ended courses effectively compared to non-open-ended courses in higher education, from an instructor perspective?

We organize our analysis into three parts, consistent with the survey structure. We first present findings related to instructor-student interaction and peer interaction in a course. Then, we discuss results around the use of assignments and exams in a course. Lastly, we show findings around the process of providing individual feedback in a course.

Table 3. Differences in the use of course approaches, including Lecture, Discussion/Seminar, Lecture only, Lecture and Discussion/Seminar, and Discussion only, between open-ended (OE) and non-open-ended (NOE) courses that are significant using a Two-proportion z-test. OE courses used discussion more often and lecture less often significantly than NOE courses. Significance levels are denoted by: *** = p-value < 0.001, ** = p-value < 0.01, * = p-value < 0.05, ^= p-value < 0.1.

Type	Lecture*	Discussion**	Lecture Only*	Lecture + Discussion*	Discussion Only^
Open-Ended	78.69%	57.83%	31.15%	36.07%	11.48%
Non-Open-Ended	91.67%	33.33%	48.33%	18.33%	5.00%

4.1.1 Instructor-student Interaction and Peer Interaction. Based on how open-ended and non-open-ended courses were operationalized, our findings suggest a significantly higher amount of interaction in OE courses than NOE courses, suggesting that learning requires richer interaction in OE courses than in NOE courses. As shown in Figure 1, discussion sessions typically involved more interaction than lectures. Based on a Two-proportion z-test, Table 3 shows that OE courses utilized lectures less often (p-val < 0.05) and discussions more frequently (p-val < 0.01) than NOE courses, indicating an emphasis on fostering bidirectional exchange of ideas in open-ended courses. We further compared the interaction time in class using a Student t-test. In Figure 1, the t-test suggests that instructors of OE courses spent a significantly bigger fraction of time on both instructor-student interaction and peer interaction in lectures (p-val < 0.05). Lastly, we found that the Blackboard Collaborate tool, messaging tools (e.g., WeChat, WhatsApp), and social media were mentioned in OE courses as technologies that facilitate interaction but not in NOE courses, suggesting the importance and need for reciprocal communication in learning open-ended material.

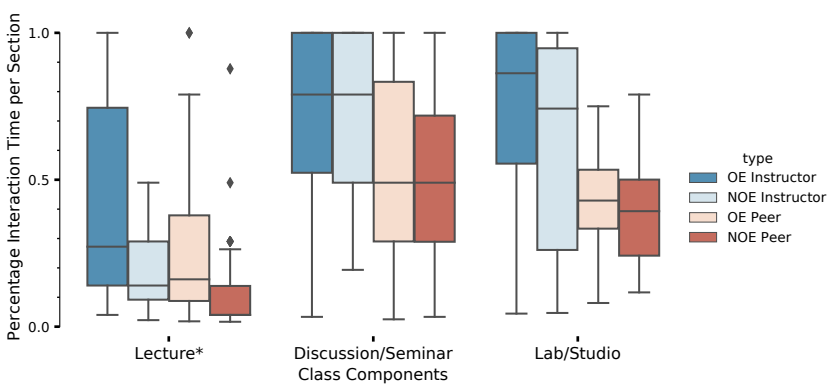

Fig. 1. A comparison of percentage of time spent on instructor-student interaction and peer interaction between open-ended (OE) and non-open-ended (NOE) courses per Lecture section, Discussion/Seminar section, and Lab/Studio section, from the instructor survey. A Student t-test was used to check if the difference in usage is significant. OE courses spent a significantly higher percentage of time on instructor-student interaction and peer interaction in lectures than NOE courses. Significance levels are denoted by: *** = p-value < 0.001, ** = p-value < 0.01, * = p-value < 0.05, ^= p-value < 0.1.

Furthermore, we solicited instructors' opinions in the survey about the degree of helpfulness and necessity of having interactions in open-ended and non-open-ended courses through a 5-point

Likert scale. The Likert scale ratings were labeled as follows – 1: Not at all helpful, 2: Somewhat helpful, 3: Helpful but not necessary, 4: Helpful and necessary, 5: Extremely helpful and necessary. As shown in Figure 2, instructors of open-ended courses rated both instructor and peer interaction significantly more valuable and necessary than instructors of non-open-ended courses in lectures and discussions. In a separate survey question where we asked instructors to choose the factor that contributed the most to students' learning outcomes among the six given major factors [3], the percentage of instructors of OE courses that ranked instructor-student interaction as the most critical factor was significantly higher than that of instructors of NOE courses (p-value < 0.05). "Instructor-student interaction" was the second most-selected factor among instructors of OE courses, while "presenting information well to the students" was second for instructors of NOE courses. "Working through a well-designed assignment" ranked first in both cases. Instructors' opinions supported our findings of richer interactions in open-ended courses based on the comparison of instructional methods in the previous paragraph.

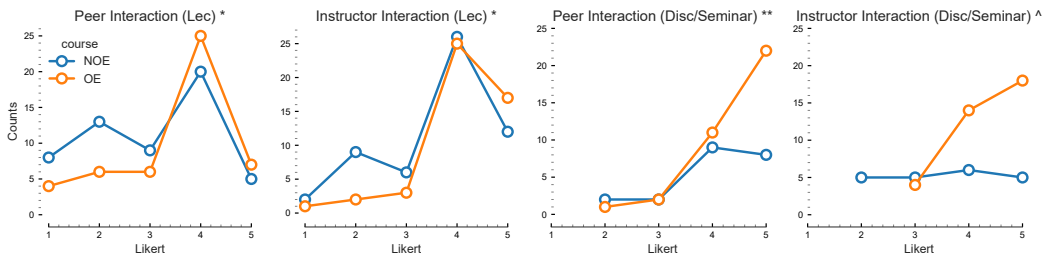

Fig. 2. A comparison of the distribution of instructor's attitudes towards peer and instructor interactions in Lectures and Discussion/Seminar between open-ended (OE) and non-open-ended (NOE) courses on a 5-point Likert scale, from the instructor survey. The Likert scale's ratings are labeled – 1: Not at all helpful, 2: Somewhat helpful, 3: Helpful but not necessary, 4: Helpful and necessary, 5: Extremely helpful and necessary. Mann Whitney U-Test is used to check if the difference in attitudes is significant. Instructors of OE courses considered peer and instructor interactions as more helpful and necessary than those of NOE courses, in both lectures and discussion/seminar. Significance levels are denoted by: *** = p-value < 0.001, ** = p-value < 0.01, * = p-value < 0.05, ^= p-value < 0.1.

We discovered three unique characteristics of interaction in open-ended courses. First, interaction in open-ended courses involved more expressions of individuality. As shown in Figure 4, there was significantly more planned one-on-one interaction with instructors in open-ended lectures (p-val < 0.05). For peer interaction, the more frequent use of peer group activity in open-ended discussion sessions (p-value < 0.05) in Figure 3 suggests that instructors in open-ended courses used a more intimate discussion environment to offer students more opportunities to engage with each other. Also, open-ended studios used class-wide activities to facilitate the sharing of individual ideas with peers more often, though this was not found to be significant due to the small sample size (60% vs. 35%).

Instructors of open-ended courses confirmed the importance of individuality in instruction via their attitudes ratings in the survey. In Figure 5, the median Likert rating among instructors of open-ended courses was "4: helpful and necessary" for both personalized instruction, and knowing

[3] 1. having information well-presented to the students; 2. interaction between instructors and students; 3. interaction among peers; 4. working through a well-designed assignment; 5. feedback to students' work; 6. effort spent by the students on their own

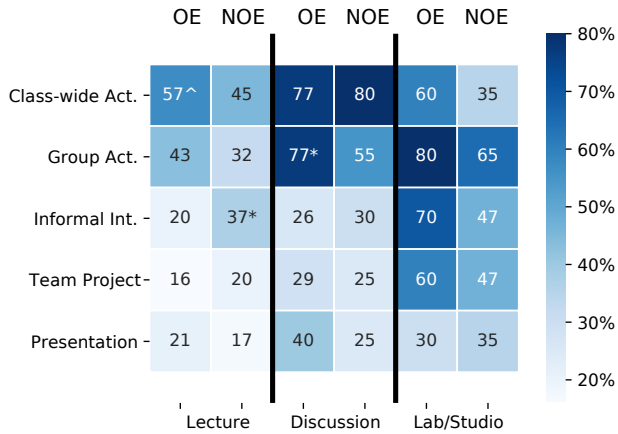

Fig. 3. A comparison of the usage of each peer interaction channel based on the percentage of courses that utilized a channel between open-ended (OE) and non-open-ended (NOE) courses in a Lecture section, Discussion/Seminar section, and Lab/Studio section, from the instructor survey. The five peer interaction channels included: class-wide activity, group activity, informal interaction, team project, and presentation. A Two-proportion z-test was used to check if the difference in usage was significant. There were significantly more class-wide activities in lectures in OE than NOE courses. However, NOE courses had significantly more informal interaction than OE courses in lectures. In discussion sections, OE courses had significantly more group activities than NOE courses. Significance levels are denoted by: *** = p-value < 0.001, ** = p-value < 0.01, * = p-value < 0.05, ^= p-value < 0.1.

students' background and preparedness on an individual basis. In contrast, the median Likert rating on the same items among instructors of non-open-ended courses was "3: helpful but not necessary", significantly lower than that of instructors from open-ended courses (p-value < 0.05). One reason for the difference could be that given that most open-ended content is value-laden, every student has their interpretation, and instructors cannot use a one-size-fits-all approach; personalization would support different student interpretations.

Second, peer interaction in OE courses was more formal, structured, and extensive than in NOE courses. In Figure 3, instructors of OE courses conducted class-wide activities (p-val < 0.1) and group activities more in lectures, while NOE lectures relied on over-the-shoulder (informal) interaction more (p-val < 0.05).

Lastly, informal and ad-hoc interaction among peers and between the instructor and students played a major role in the studio and lab sessions of OE courses, but not so much in OE courses' lecture or discussion sessions. For instructor-student interaction, Figure 4 shows that 70% of the open-ended lab/studio sessions involved ad-hoc one-on-one interaction, a contrast from the 18% and 23% in open-ended lectures and discussion/seminar sessions, respectively. Similarly, for peer interaction in Figure 3, 70% of open-ended lab/studio sessions utilized informal interaction but only 20% and 26% open-ended lectures and discussion/seminar, respectively, included informal interaction. The differences suggest that ad-hoc interaction is a unique and frequently-used interaction channel in open-ended labs and studios.

In conclusion, from the instructor's perspective, reciprocal instructor-student interaction and peer interaction were more valuable and necessary for the teaching and learning experiences in open-ended courses than that in non-open-ended courses. They incorporated individuality significantly more in open-ended courses, making it a necessity instead of a recommendation. While the interaction was more structured and extensive in open-ended lectures and discussions/seminars,

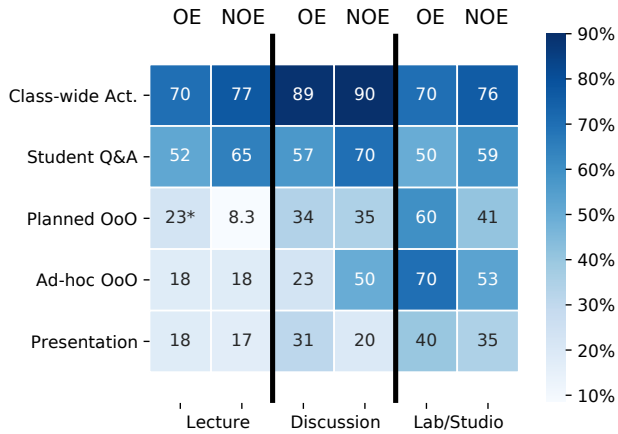

Fig. 4. A comparison of the usage of each instructor-student interaction channel based on the percentage of courses that utilized a channel between OE and NOE courses in a Lecture section, Discussion/Seminar section, and Lab/Studio section, from the instructor survey. The five instructor-student interaction channels included: class-wide activity, student-initiated Q&A, instructor-planned one-on-one, ad-hoc one-on-one, presentation. A Two-proportion z-test was used to check if the difference in usage was significant. There was significantly more planned one-on-one interaction between the instructor and students in OE courses than in NOE courses. Significance levels are denoted by: *** = p-value < 0.001, ** = p-value < 0.01, * = p-value < 0.05, ^= p-value < 0.1.

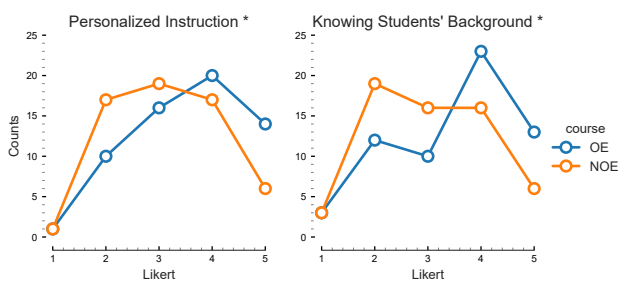

Fig. 5. A comparison of the distribution of instructors' attitudes towards personalized instruction and knowledge of students' backgrounds between open-ended (OE) and non-open-ended (NOE) courses on a 5-point Likert scale, from the instructor survey. The Likert scale's ratings are labeled – 1: Not at all helpful, 2: Somewhat helpful, 3: Helpful but not necessary, 4: Helpful and necessary, 5: Extremely helpful and necessary. The Mann Whitney U-Test was used to check if the difference in attitudes was significant. Instructors of OE courses tended to rate personalized instruction as helpful and necessary while more NOE instructors rated it as good-to-have or only somewhat helpful. Similarly, instructors of OE courses found knowing students' background to be more necessary and helpful than those of NOE courses. Significance levels are denoted by: *** = p-value < 0.001, ** = p-value < 0.01, * = p-value < 0.05, ^= p-value < 0.1.

over-the-shoulder interaction played an indispensable role in open-ended labs/studios. These findings make up the first part of our answer to RQ 1.

4.1.2 Types of Assignments and Exams. In the instructor survey, we found that while assignments and assessments were considered the most-valuable to learning in both open-ended and non-open-ended courses, they were more open and diverse in open-ended courses.

From the perspective of instructors of both open-ended and non-open-ended courses, working through a well-designed assignment mattered the most to students' learning outcomes, as it ranked first in both groups among the six major factors[4] that contribute to students' learning. This ranking justifies the criticality of this pedagogical element.

In order to define types of assignments, we created a preliminary coding of assignments presented in the syllabus of courses across two departments in each of the subject areas of Natural Sciences, Mathematics, Engineering, Social Sciences, Humanities, and Arts. Based on the codes, we defined four types of assignments based on their openness and goals.

- **Close-ended solution-oriented assignments:** finding solutions to a well-defined problem with a definite answer, e.g. matrix multiplication, finding the compounded interest rate of a bond.
- **Open-ended solution-oriented assignments:** finding solutions to a well-defined problem with indefinite answers, e.g. a software solution that classifies pictures of dogs and cats, describing means to tackle climate change
- **Open-ended argument-oriented assignments:** providing arguments to a topic or issue, e.g. a reflection on a reading, a paper on the definition of consciousness
- **Open-ended artifact creation:** creating open-ended artifacts, e.g. writing a piece of literature, performing a play, creating a piece of artwork

In case we missed other types of assignments, we included an "Other" option in the survey. One participant selected "Other" as their only option. We determined their provided description, "writing news articles according to specified [sic] format," fell into the open-ended artifact creation assignment category.

In Figure 6, most open-ended courses included at least some form of open-ended assignment, whereas non-open-ended courses mostly had close-ended assignments. Instructors in 83.3% of the NOE courses used close-ended assignments, while only those in 24.6% of the OE courses incorporated them (p-val < 0.001). Instructors in open-ended courses offered a more diverse variety of assignments. OE courses incorporated a heavy-use of open-ended argument-oriented assignments (p-val < 0.001) and open-ended artifact creation (p-val < 0.01), unlike NOE courses.

Open-ended course exams included open-ended questions almost all the time based on Figure 6 (p-val < 0.001). Furthermore, open-ended courses incorporated exams significantly less often (p-val < 0.001). Reasons instructors did not utilize exams in open-ended courses included having open-ended content that was difficult to test through an exam, needing iterations on a deliverable, involving content based on personal experience, and evaluating performance based on the student's progress.

The openness of the assignment and assessment methods used in OE courses often required the accompaniment of a flexible grading schema. As a result, there was an increased workload and difficulty in grading. The average time instructors of open-ended courses took to grade one assignment submission was 18.14 minutes, about 60% longer than that of a non-open-ended course instructor, which was 11.25 minutes.

Overall, as a part of the response to RQ 1, the type of assignments and assessments used in a course was another factor that differed in the teaching and learning experiences of open-ended

[4] 1. having information well-presented to the students; 2. interaction between instructors and students; 3. interaction among peers; 4. working through a well-designed assignment; 5. feedback to students' work; 6. effort spent by the students on their own

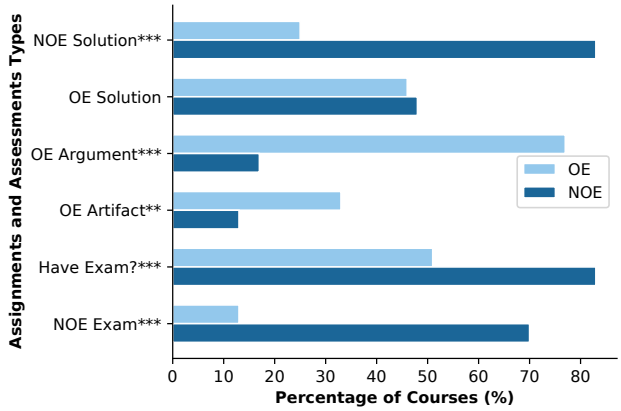

Fig. 6. A comparison of the usage of various types of assignments and exams based on the percentage of courses that utilized them in open-ended (OE) and non-open-ended (NOE) courses, from the instructor survey. The four types of assignments included: non-open-ended (NOE) solution-oriented assignments, open-ended (OE) solution-oriented assignments, open-ended (OE) argument-oriented assignments, and open-ended (OE) artifact creation. The two types of exams included exams with only non-open-ended (NOE) problems and those with some open-ended (OE) problems. The "Have Exam?" row refers to the percentage of courses that had an exam. A Two-proportion z-test was used to check if the difference in usage was significant. NOE courses included NOE solution-oriented assignments and exams (especially NOE exams) significantly more than OE courses. OE courses utilized OE argument-oriented and OE artifact-creation assignments significantly more than NOE courses. Significance levels are denoted by: *** = p-value < 0.001, ** = p-value < 0.01, * = p-value < 0.05, ^= p-value < 0.1.

courses and non-open-ended courses. Both assignments and exams in open-ended courses often had more than a small number of correct answers, while those in non-open-ended courses did not. Types of assignments were more diverse in open-ended courses, and included solution-oriented, argument-oriented, and artifact creation assignments. Assessment methods were also more flexible, considering that an exam was not a very popular option in open-ended courses.

Table 4. A comparison of whether individual feedback was provided and when it was provided in open-ended (OE) and non-open-ended (NOE) courses via a Two-proportion z-test, from the instructor survey. Open-ended courses provided significantly more individual feedback than non-open-ended courses. Significance levels are denoted by: *** = p-value < 0.001, ** = p-value < 0.01, * = p-value < 0.05, ^= p-value < 0.1.

Type	In Progress*	After Submission***	Answer Key Only***	No Feedback*
Open-Ended	60.66%	93.44%	0.00%	0.00%
Non Open-Ended	45.00%	73.33%	16.67%	5.00%

4.1.3 Individual Feedback. Our findings suggest that individual feedback played a larger role in open-ended courses than in non-open-ended courses. Overall, students in open-ended courses received significantly more individual feedback, both after assignment submission (p-val < 0.001)

and while the assignment was in progress (p-val < 0.05), as shown in Table 4. Students in all of the open-ended courses received individual feedback at some point, while this was not the case for non-open-ended courses.

When we asked the instructors in the survey to rate how much they valued individual feedback in their course on a 5-point Likert scale, instructors of open-ended courses rated it significantly higher (p-value < 0.05), with a median of "4: helpful and necessary", as shown in Figure 7. A possible reason for such a difference is that, unlike in NOE courses where feedback may not be necessary for students who already arrive at the correct answer on their own, in OE courses, even the best students could benefit from feedback because there is no single correct answer and there are always things to improve on.

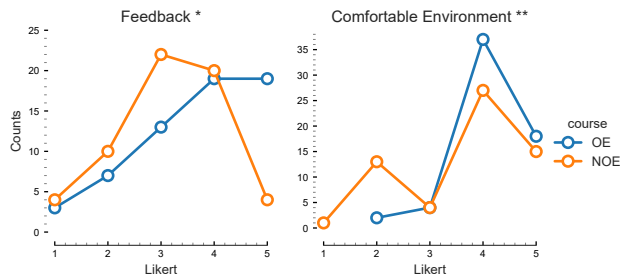

Fig. 7. A comparison of the distribution of instructors' attitudes towards the value of individualized feedback and having students be comfortable around each other in open-ended (OE) and non-open-ended (NOE) courses on a 5-point Likert scale, from the instructor survey. The Likert scale ratings were labeled – 1: Not at all helpful, 2: Somewhat helpful, 3: Helpful but not necessary, 4: Helpful and necessary, 5: Extremely helpful and necessary. The Mann Whitney U-Test was used to check if the difference in attitudes was significant. Instructors of OE courses rated both individualized feedback and making students feel comfortable with each other as more valuable and necessary than those of NOE courses. Significance levels are denoted by: *** = p-value < 0.001, ** = p-value < 0.01, * = p-value < 0.05, ^= p-value < 0.1.

In terms of the channels that provided individual feedback, there was—on average—a wider variety of channels (2.46 channels per course) in OE courses than in NOE courses (1.62 channels per course). When instructors of open-ended courses provided feedback, they provided verbal feedback (p-val < 0.05) and written text feedback on paper (p-val < 0.05) significantly more often than instructors in NOE courses (Figure 8). A possible explanation for the frequent use of verbal feedback is because it is a form of in-the-moment feedback that is the norm in a studio environment. Based on our findings in Section 4.1.1, in-the-moment interaction was an essential interaction channel in studios of open-ended courses.

Besides instructor feedback, instructors of open-ended courses made use of peer feedback via various channels significantly more than instructors of non-open-ended courses. Instructors rated the quality of peer feedback as acceptable with appropriate scaffolding and learning, as shown in Figure 8. In addition, instructors of open-ended courses rated the need for having students feel comfortable around each other significantly higher than instructors of non-open-ended courses (p-val < 0.01) in the survey, as shown in Figure 7. The mode of the attitude distribution among instructors of open-ended courses was between "5: extremely helpful and necessary" and "4: helpful and necessary," while it was "3: helpful but not necessary" among instructors in non-open-ended courses. One possible explanation may be the more frequent use of peer feedback in open-ended

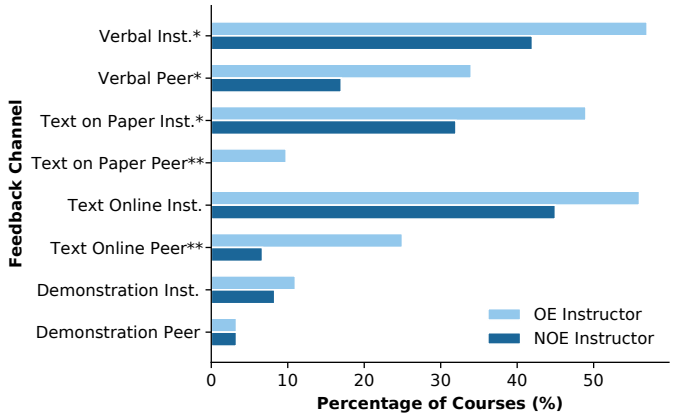

Fig. 8. A comparison of the usage of various instructor and peer feedback channels in OE and NOE courses based on the percentage of feedback-providing courses that utilized a channel, from the instructor survey. The eight feedback provision channels included: feedback via verbal comments, writing on paper, text online, and demonstration, from instructors and peers respectively. A Two-proportion z-test was used to check if the difference in usage was significant. OE courses provided feedback via five of the eight feedback channels significantly more often than NOE courses, all except instructor feedback via online text and instructor and peer feedback by demonstration. Significance levels are denoted by: *** = p-value < 0.001, ** = p-value < 0.01, * = p-value < 0.05, ^= p-value < 0.1

courses and the value-laden nature of open-ended assignments. To conduct both of these learning activities effectively, a class requires a high level of trust [12].

In conclusion, providing individual feedback was a more crucial and necessary pedagogical element in open-ended courses than in non-open-ended courses. Instructors of open-ended courses provided both formative (in-progress) and summative (after submission) feedback more often than those of non-open-ended courses via a wider variety of channels. Peer feedback of all forms was more prevalent in open-ended courses but required scaffolding. There was also a greater need for students to feel comfortable around each other.

4.2 The Student Perspective (RQ 2)

In the previous section, we identified what mattered to instructors in teaching OE courses. In this section, we explore the students' viewpoints to understand points of alignment and to answer our second research question:

RQ 2 From a student perspective, what are student preferences for how instructors should teach open-ended courses?

We begin by presenting the students' views on instructor-student interaction and peer interaction. Then, we discuss findings related to assignments and learning activities in a course. Finally, we describe students' experiences surrounding the process of receiving individual feedback.

4.2.1 Instructor-student Interaction and Peer Interaction. Based on the student survey, an open-ended course that students enjoyed typically had an above-average amount of instructor-student interaction and peer interaction compared to other open-ended courses. When students listed the reasons why a course was their favorite, based on our thematic analysis, the ability to exchange ideas via discussion or via looking at others' work ranked third among the reasons why an OE

course was a student's favorite (Table 5). As to why a student did not enjoy an OE course, lack of discussion ranked first. These findings are in line with the instructors' views on interaction needs.

Instructor-student interaction ranked first among the students' crucial factors [5] for learning. When we solicited students' preferred degrees of these interactions compared to their favorite OE course, 38.18% wanted more interaction with instructors; no one wanted less. For peer interaction, 27.27% wanted more; 5.5% preferred less. From this, we infer that students prefer more instructor-student interaction, while they favor sufficient but not an overwhelming amount of peer interaction. These findings align with the instructors' opinions for the most part but add a more nuanced understanding of the need for peer interaction.

Table 5. Top eight student reasons (identified with thematic analysis), along with frequency of mentions, for favoring and for disliking an open-ended course.

Reason for Favoring	**# Mentions**	**Reason for Disliking**	**# Mentions**
Open-ended tasks that encourage self-expression	17	Lack of discussion	9
Thought-provoking and require introspection	16	Too difficult or too much work	8
Idea exchanges via discussion or learning from others' work	12	Uninterested content	8
Personal interest	6	Not open to subjective opinion or too structured	7
Useful and informative	6	Unclear instruction	7
Engaging instructor	5	Unfriendly or judgemental peer critique	3
Fun and interesting	5	Lack of instructor' responses	3
Welcoming atmosphere	4	Over-populated class	3

4.2.2 Encourage and Respect Open-ended Self-expression. Students preferred an open-ended course when it encouraged free self-expression, promoted introspection, and asked thought-provoking questions via open-ended tasks and assignments. These two reasons ranked first and second, respectively, among the reasons students enjoyed an open-ended course. The rankings suggest that a well-designed assignment or learning activity should be thought-provoking and require students to introspect and create their own solutions or artifacts. As a result, the ideas, opinions, and artifacts created in an open-ended course can be highly personal and value-laden.

Students mentioned two things that they appreciated in how these assignments or learning activities were evaluated. First, students preferred evaluation criteria that did not limit what was considered right or wrong. Not being open to subjective opinion or being too structured ranked 4th among the reasons students disliked an open-ended course (Table 5). When the course encouraged self-expression and creativity, students reported that they focused more on the process instead of getting the right result. Students' opinions were in line with the finding from the instructor survey that open-ended courses should design assignments that are open-ended in their evaluation standard in order to give students enough room to express themselves freely.

[5]Factors to choose from included: lecture by the instructor, interaction with peers, interaction with the instructor, feedback or critique to the assignments, and self-studying.

Second, students preferred a welcoming and unbiased atmosphere for evaluation. Having a welcoming atmosphere ranked 8th for favoring an open-ended course while having an unfriendly or judgemental peer critique environment ranked 6th among reasons for why students disliked an open-ended course (Table 5). Students reported that with a welcoming evaluation environment, they felt more comfortable with putting forward their ideas, opinions, and artifacts. This finding coincides with instructors' opinions on the need to have students comfortable around each other in open-ended courses in the instructor survey.

4.2.3 Individual Feedback. Students valued individual feedback in open-ended courses. In their favorite open-ended courses, 69% of the students considered the feedback they received for their assignments to be very instructive or extremely instructive, and 27% rated it as somewhat instructive. 20% of the students selected receiving feedback to assignments as the most essential factor for learning compared to the four other factors, including attending lectures, interacting with instructors, interacting with peers, and self-studying. On average, each course utilized approximately 2.5 distinct feedback channels concurrently. The above findings from the student survey reaffirm the instructor survey's conclusion that it is critical for open-ended courses to provide extensive feedback via multiple channels.

In the instructor survey, we found that peer critique was used more frequently in open-ended courses than in non-open-ended courses. Similarly, in the student survey, 55% of students' favorite open-ended courses adopted peer assessment. Regarding the effectiveness of peer assessment as a means of peer interaction, students ranked it as the seventh most effective out of ten channels. Experiencing judgemental peer assessment was also a reason why students disliked a course. These results suggest that instructors consider the limitations of peer assessment, provide sufficient scaffolding, and not rely on it as the sole approach to provide feedback or facilitate peer interaction.

4.3 The Five Elements of Successful Open-ended Courses

So far, in Section 4.1 and Section 4.2, we presented findings around how to teach an open-ended course well based on how instructors have been teaching, what instructors think, and what students prefer. Synthesizing the instructor and student surveys' findings, we summarize them into the five elements below that differentiated the teaching and learning experiences in open-ended courses from that in non-open-ended courses. These elements were either not valued as necessary, or not considered as applicable to non-open-ended courses. With a clear view of critical elements of open-ended courses that make for a positive experience, we can better capture the challenges and opportunities of incorporating these elements into open-ended courses online and at scale.

(1) Include an extensive amount of instructor-student interaction and peer interaction in various forms via multiple channels for idea exchange and knowledge sharing; instructor-student interaction is of higher priority and a critical factor to student learning *[Section 4.1.1 & Section 4.2.1]*
(2) Utilize well-designed open-ended assignments and consider flexible assessment and evaluation methods to allow self-expression *[Section 4.1.2 & Section 4.2.2]*
(3) Provide customized individual feedback from instructors and peers through a variety of feedback channels to open-ended work for all students; instructors should be wary of over-reliance on peer feedback and should use it with scaffolding *[Section 4.1.3 & Section 4.2.3]*
(4) Foster an open and unbiased environment for learning and evaluation while encouraging introspection, self-expression, and creativity *[Section 4.1.3 & Section 4.2.2]*
(5) Incorporate students' individual backgrounds into personalized instruction as a requirement rather than as a possibility *[Section 4.1.1]*

5 INTERVIEW RESULTS—CHALLENGES AND OPPORTUNITIES OF TEACHING OPEN-ENDED COURSES ONLINE AT SCALE (RQ 3)

In the previous section, we derived pedagogy criteria that differentiated the teaching and learning experiences in an open-ended course from that in a non-open-ended course. To better understand what is needed to incorporate these criteria into open-ended courses online at scale, in this section, we present the findings to our final research question:

RQ 3 What are the challenges and opportunities in teaching university-level, open-ended courses effectively, online at scale in a university setting?

To answer this question, we interviewed 11 instructors about the challenges and opportunities they have faced or would face when teaching open-ended courses with a large number of students or online in a university setting. The interview results in this section answer our third research question. We organize the results into five subsections, based on the five critical pedagogy elements specific to open-ended courses summarized in Section 4.3.

The instructor discussed a total of five small in-person courses, five large in-person courses, five large online courses, and one small online course in the interviews. Instructors self-labeled small and large classes by comparing their class to the perceived department class size averages. For a more detailed description of the interview sample, please refer to Section 3.3.

5.1 Element 1: Extensive Interaction

> *It's all about conversation. … So, then the core question about scalability is about conversation. [P8]*

A graphic designer stated the quote above as the biggest challenge to their course. Overall, six instructors found it harder to co-construct knowledge with students with a lecture as the preferred way to share information in a large class.

Four instructors mentioned having a less discussion-friendly or studio-friendly classroom environment when teaching in a large lecture hall. It prevented students from having back-and-forth discussions that were audible to the entire class, performing group activities, or engaging physically using a big roll of white paper and post-it notes. As we have seen in the previous section, these interaction activities played a critical role in OE courses compared to NOE courses. As courses get large, time also becomes a limiting factor. Instructors had to control the time and flow of the discussions by having a more constrained discussion that limited conversations among students. They mentioned not having enough time to cover every student in a discussion. They considered this as a challenge to satisfy the requirement of giving students sufficient opportunities to participate and express themselves in an open-ended course.

Despite the above challenges of having insufficient interaction in large scale open-ended courses, four instructors mentioned the benefit of increased diversity in opinions and ideas for open-ended content in a large class due to a more diverse student body. Unlike non-open-ended courses where most problems only have one or a few correct solutions, open-ended courses focus on open-ended problems that have a positive marginal benefit as student population increases. A studio art instructor said the following:

> *That's beneficial. There would be more kinds of answers to assignments. There would be different perspectives that would come up in critique. [P1]*

Teaching online presents another set of challenges and opportunities for facilitating interaction. We will first present challenges with synchronous online interaction. Overall, five interviewees questioned the quality and scalability of synchronous interaction online. Instructors considered interaction in conversations via video or audio conferencing to be not as effective due to missing

visual cues. Instructors were concerned about students' unequal access to high-speed internet that led to signal loss or inability to participate. Our survey found that, for both open-ended and non-open-ended courses, online sessions with 35 students or more did not utilize any synchronous interaction channel, which is in line with the concern of its scalability shown in the interviews.

Next, we will move on to the challenges with asynchronous forum discussion, the most commonly used interaction channel in online OE courses based on our survey. Three instructors commented that such asynchronous discussion exhibited artificial flow and produced shallower conversations. Because of the delays in responses to posts, students often did not return to the conversation after replying to the initial post. This issue stood out more in open-ended courses because discussions in open-ended courses were commonly conversations with multiple rounds of back-and-forth, unlike the more specific question-and-answer pattern in non-open-ended courses. In addition, with the asynchronous forum, in most cases, the students that answered first essentially removed the need and opportunity for others to participate.

Three instructors also expressed concerns about the lack of real engagement in these forum interactions. A fashion design instructor described how students acted when they were asked to comment on each other's posts:

> *So it's like here, marking it off the list. Oh, you want me to do this? Okay, we'll do this. [P7]*

While instructors believed that instructor intervention could facilitate better discussions, they acknowledged that amount of time and effort needed would be overwhelming. P7 commented:

> *If I were to create more opportunities for students to create content for the online class, that would require a lot more work on my part to oversee that. [P7]*

In addition to more formal interaction, we described in the previous section that over-the-shoulder interaction was a critical part of the studio environment experience for open-ended courses. In our survey, online OE courses had significantly less ad-hoc peer interaction than in-person OE courses. An art education instructor interviewee elaborated on the lack of support for in-the-moment peer interaction in an online studio environment.

> *... it's those little things that you miss. The ability to look at the person next to you and say I like what you're doing, how do you do that? [P9]*

However, teaching an open-ended course online brings opportunities to improve interaction with a new medium. Two Philosophy instructors stated that since discussion online can be a mandatory class requirement, the students that preferred writing to speaking would have more opportunities to interact online than in-person. P7 instructor was impressed by the depth of initial posts made on a forum, saying that they were more profound than the in-person responses likely because of the extra time for deliberation available online. A philosophy lecturer proposed that having discussions online potentially allowed easy archiving and reviewing of the takeaways with an appropriate tool. These are all potential benefits we can leverage when offering open-ended courses online at scale.

5.2 Element 2: Flexible Assignments and Assessments

When teaching open-ended courses at a large scale, some instructors altered the usage of open-ended assignments. The survey suggested that each large open-ended assignment required 40 minutes to grade per student on average. The interviewees reported similar times. As a result, some instructors either reduced the number of deliverables or used a more constrained assignment with fewer possible answering directions to cope with the increased grading workload in a large class. In courses that used group projects with a larger number of students (more than 2), three instructors raised challenges about balancing everyone's role and dealing with complex logistical issues. For

example, an Architecture instructor thought students did not learn all the required materials and skills evenly in a group project because students tended to do what they were already good at in a team:

> *Generally what happens in a team is that, you just do what you're already good at. That's not the point. College, you want to be, for architecture you need to be good at everything. So, some people, they can just do really beautiful imagery, so that's all they do. But they're never learning to think about these deeper issues with society, they're not even learning how to design proper spaces. ... I know in reality, that when the students go to work, they'll have to work in a team, so they do need to learn from those experiences and learn how to negotiate and all that, but I don't think it's a solution for over-crowding or large class sizes. No, I don't. [P2]*

Two instructors mentioned that teaching online brought challenges when assignments required the creation of artifacts that relied on specific physical equipment or spaces, e.g., using a sewing machine in a fashion design course. Teaching a process such as sewing online is hard. It becomes even harder when some students may not have access to the same physical equipment or spaces. In most cases, such assignments could not be replaced by an alternative that did not have this physical requirement as students often need to learn through experience in open-ended courses. Hence, it is essential to consider how inclusive a learning at scale platform is for various types of assignments and assessments.

5.3 Element 3: Individual Open-ended Feedback

All of the instructors we interviewed considered instructor feedback an irreplaceable part of open-ended courses while admitting that it was time-consuming; this aligned with our survey findings. Six instructors brought up the time-to-quality trade-off in providing instructor feedback in a large class, for both written narrative feedback and verbal feedback. Written feedback was usually provided after class, while verbal feedback was often offered in class. It was not uncommon for the instructors to schedule a non-trivial block of time with each student or group in class to provide formative feedback in open-ended courses. A graphic design instructor considered instructor feedback an organic part of the development process of a deliverable that could not be removed or shortened:

> *I'm being part of the development. I'm not only checking the due date. Like, it's not a check list. [P3]*

Instructors also mentioned the lack of in-the-moment formative feedback in an online studio course. An art education instructor expressed their preference as:

> *It's much easier as a teacher also to walk through a studio class and to give one on one feedback verbally than it is to click and go through someone's weekly studio portfolio and try to give meaningful feedback when it's already done and when it's typing out comments rather than just being able to speak about what you're seeing or noticing in that moment. [P9]*

Besides instructor feedback, our survey results suggest that peer feedback was used more often in open-ended courses than in non-open-ended courses. While peer critique is a skill students in open-ended courses should learn, instructors we interviewed were hesitant to rely on it heavily. Two instructors said that peers were reticent to give critical feedback that helped. Three other instructors thought that peer critique showed inconsistent objectivity and a lack of context when assessing subjective content. For example, the P7 instructor commented:

> *As to the quality of their comments, I think it really varies because people come with different levels of being able to hold themselves to standards other than their own personal likes or dislikes. [P7]*

Overall, it is unclear how to provide sufficient formative and summative individual feedback from instructors in open-ended courses at scale. Instructors were also doubtful about the quality of peer feedback.

5.4 Element 4: A Welcoming Environment for Self-expression

> *The small class is intensely personal. They are hanging work that came from their soul, twice a week. And that's really threatening ... my students and I trust each other ... Seldom will that happen in the big class. [P6]*

As an advertising instructor pointed out in the above quote, trust is critical in an OE course; this is consistent with what we found in the instructor and student surveys. Instructors expressed concern about the diminished trust in large and online courses. Four instructors stated that increased anonymity and decreased interaction amount in a large class resulted in fewer personal discussions in their courses. Three instructors found it challenging to build trust, break formalities, and build a community in an online setting. Two instructors further stated that large class sizes often resulted in severe evaluation apprehension and made the class environment less friendly. An Art Education instructor described what happened in their large class:

> *There's [sic] like 10 students that speak up. I'm guessing that there's [sic] more questions than that and not everybody feels comfortable raising their hand in a large format class. I know as a student I probably wouldn't have. [P9]*

Another concern with self-expressing online was the lack of emotional exchange. Since emotion comes naturally as a part of self-expression in open-ended courses, but not so much with a standard answer in non-open-ended courses, P6 pointed out their worry of lack of emotional exchange in online teaching:

> *... the technology is a wall to protect from emotions. And emotions are irrevocably attached to learning. If I can make a point in an emotional way that they understand and they're touched, they'll remember it. If I do it talking online, through a computer screen. Pfft, gone. [P6]*

On a positive note, P7 did see hope in overcoming these challenges:

> *I was very surprised by the level of intimacy and trust that a lot of students have shown in the online framework, that I have learned some things about them. [P7]*

Characteristics of the online medium provided some opportunities to enhance the level of comfort in the class environment. Instructors appreciated the chance of convenient reflection and correction on errors in online discussions, both logistically and psychologically, compared to in-person discussions. An online course could also address complaints about limited privacy in a small in-person studio classroom by managing various levels of access online.

5.5 Element 5: Personalized Instruction

All interviewees utilized some form of personalized instruction in their small open-ended courses, e.g., adapting instructional style and evaluation standards based on students' backgrounds. Large class sizes undermined the students' individuality in the instruction of open-ended courses. Four instructors mentioned their limited attention when interacting with every student one-on-one in a large class. Unlike need-based interaction in NOE courses, both students that excel and those that struggle need interaction time with the instructor. P2 described a dilemma in their larger class:

> *Because I'll spend, as a teacher you feel like you need to give assistance to the students who are struggling or aren't getting it. But the reality is, is that if you look at how competitive architecture is, it's the good students who should be getting the most attention, because they have a really good future ahead of them. [P2]*

Three instructors worried about the degree of specificity of their feedback in a large class since they did not know the students on an individual basis. P7 described the effort they put in to personalize each feedback response; this is not scalable:

> *I think probably too much about how the person might receive what I'm writing. I really try to think about who they are and writing specifically sometimes differently, especially international students. I'm really trying to be as clear and helpful as possible. [P7]*

Having fewer personal connections with students not only impacted specificity of instructors' feedback, but also how rewarding the teaching experience was for an instructor P5:

> *... just the amount that I know students, it's much more limited. Part of that's just less fun for me and for them. Like it feels impersonal in a certain way. ... So that has a huge effect on like how rewarding it is to teach... [P5]*

Overall, time and effort were common themes across the interviews. Instructors reported having to prepare more thoroughly for a large class to keep students from various backgrounds engaged. Given their immense effort, many still found personalized instruction challenging to accomplish in a large class, and sometimes even in a small class.

5.6 Summary of Interview Findings

So far, we have presented the challenges and opportunities of supporting an open-ended course well online at scale in a university setting in detail. Below is a summary of our response to RQ 3 in this section.

- **Sufficient and In-depth Interaction:** The major challenge of having sufficient interaction in a large class is the scalability of conversations with limited time and in non-discussion-friendly or non-studio-friendly environment. In an online class, it is challenging to maintain the quality and scalability of synchronous interaction online, as well as the flow, depth, and level of engagement of conversations via asynchronous interaction online. Over-the-shoulder interaction is also missing online.
  Increased diversity in peer-generated content is an opportunity for teaching a large class. Opportunities for hosting interaction online include having more deliberation time, being more friendly to students who prefer writing to speaking and creating a searchable archive for conversations.
- **Flexible Assignments and Assessments:** In a large class, it is challenging to scale an open-ended assignment due to an increased grading workload. Instructors had concerns about using an increased group size to resolve this challenge. In an online course, there are challenges in incorporating artifact creation that requires specific physical equipment or spaces.
- **Extensive Open-ended Feedback:** When teaching a large class, instructors faced the challenge of providing individual feedback to students as an organic part of the development process of a deliverable. Another challenge for the instructors is ensuring the quality of peer feedback.
- **A Welcoming Environment for Self-expression:** A major challenge for maintaining the level of self-expression in a large class is the increased anonymity, diminished trust, and severe evaluation apprehension that leads to a less welcoming environment. Challenges for

encouraging self-expression online include difficulty in breaking formalities, fostering a sense of community, and enabling emotional exchanges.
From another perspective, teaching online could encourage self-expression by offering an environment that can maintain an appropriate level of privacy and be friendly to self-reflection and error correction.

- **Personalized Instruction:** Instructors were challenged to offer personalized instruction, interaction, and feedback to students in a large class due to limited attention and loss of personal connection.

6 DISCUSSION

In this section, we aim to identify open challenges and opportunities of teaching open-ended courses online at scale in a university setting that are not yet addressed by the state of the art course design, policy, and educational technologies. Since mainstream commercial MOOC platforms are the most popular and well-studied platforms for teaching courses online at scale, we use them as a reference point for the current state of the learning-at-scale environment. For each challenge and opportunity found in the previous section, we first describe the degree to which they are currently addressed in mainstream MOOC platforms. Then, we summarize existing research that attempts to resolve the issue and point out areas that require future research.

6.1 Facilitate In-depth Conversations

While prior work has shown the importance and effectiveness of interaction in online courses [68], we found that *structured, extensive, and reciprocal* instructor-student interaction and peer interaction were more valuable and necessary in open-ended lectures and discussions than in the non-open-ended ones (Section 4.1.1). In the interviews, we learned that conversations were the primary type of interaction in open-ended courses (Section 5.1), and instructors reported difficulties creating a discussion-friendly environment both in a large class and online. From the survey findings (Section 4.1.1), we concluded that open-ended courses offered online at scale in a university setting should support in-depth conversations between instructor-student and among peers class-wide, in groups, and one-on-one. More specifically, the interaction between instructors and students was more critical than peer interaction.

Currently, most MOOC platforms utilize video or audio conferencing tools and chat rooms for group or one-on-one synchronous communication. Discussion forums, social media (e.g., Facebook, Twitter) and email communication are often used for asynchronous interaction class-wide, in groups and one-on-one [1, 40]. Prior research found that peer interaction was the dominant type of interaction on MOOCs, whereas there was little instructor-student interaction [43].

Compared to asynchronous communication mediums, guided synchronous video conferencing calls provide more natural discussion flow and better engagement, and thus could support small group in-depth conversations well [46]. However, as found in our interviews, they were not accessible to students without high-speed internet. Another open challenge derived from our interviews was the scalability of these calls. If the calls occurred among a small group of students, then the challenge became the scalability of instructor-student interaction across many groups, given instructors' limited time and bandwidth.

Asynchronous interaction channels, such as discussion forums, had issues with unnatural flow and a lack of real engagement, impacting the frequency of turn-taking and the depth of discussion, as found in the interviews. Existing work has uncovered similar issues in MOOCs—a lack of substantial discussion in chat rooms [16], discussion forums, and social media [19, 1]. Limited work addresses these issues via improved MOOC platforms. Coetzee et al. [16] found that having a more visible and pervasive chat interface encouraged significantly more substantive interaction. Yang

et al. [83] developed a discussion thread recommendation framework that tried to lead students to respond to the most relevant threads.

There are a few studies in online distance education that have looked into the above issues from an instructional design perspective. Christopher et al. [15] suggested that the use of an evaluation rubric may encourage higher-level thinking in discussion forums. Dennen* [22] found that guidelines, deadlines, feedback, and instructor presence generated better conversations in discussion forums, which is consistent with what instructors said in our interviews. Future open-ended courses offered online at scale in a university setting could consider integrating these instructional designs.

While both instructors and students considered instructor interaction to be more critical than peer interaction in the surveys, current open-ended MOOCs employ little of it [43], relying mostly on volunteer mentors to assist the course instructors. Existing work on how to increase instructor interaction in learning-at-scale environment is limited. Future advances in algorithms or human-in-the-loop designs may help instructors better prioritize students' needs. As an example, Chandrasekaran et al. [13] trained a binary-classification model to identify discussion forum posts that required an instructor's intervention.

Overall, we suggest that future work adds to the existing discussion channels or creates new and necessary channels to better support and incentivize in-depth conversations class-wide, in group and one-on-one settings. Future studies could also develop algorithms to effectively increase instructors' involvement without overwhelming the instructors. Technology alone will not solve many of these problems. For example, for instructors to increase one-on-one interactions with students, more time and more instructors would help. Understanding the time required for instructors to effectively teach large online courses would help university policy-makers reflect on and re-evaluate teaching policies and expectations. By doing so, this would benefit both students and instructors.

6.2 Create A Studio-friendly Environment

The studio is an indispensable component of many open-ended courses, especially in Art & Design [55]. Our surveys and interviews suggested three key characteristics of studio sessions, over-the-shoulder peer learning (Section 4.1.1, Section 5.1), in-the-moment instructor feedback (Section 4.1.1, Section 5.3), and peer learning via work-sharing and discussion (Section 4.1.1, Section 4.2.1).

Instructors in our interviews reported the complete loss of over-the-shoulder peer learning and in-the-moment instructor feedback in an online course. Current MOOC platforms support neither of them.

As its name suggests, "in-the-moment" feedback in a large online course is only possible through synchronous channels. Synchronous channels with visual cues, i.e., video conferencing, suffer from limited accessibility, as mentioned above. While being "in-the-moment" is extremely challenging in this context, future work may design innovative *asynchronous* solutions that achieve a similar goal. For example, PeerStudio provides asynchronous in-progress peer feedback [47]. While feedback receivers can request and receive feedback on in-progress work via PeerStudio, in a studio, in-the-moment feedback is sometimes initiated by the feedback provider based on the needs and context they observe. Future research could develop automated solutions that pick up on cues that indicate a need for feedback in an online studio environment. Instructors could also train more qualified staff members and develop course policies around providing frequent in-progress feedback in an online environment at scale.

As for the lack of over-the-shoulder peer learning, to the best of our knowledge, no existing research provides techniques to address this problem in a large scale studio environment. In the space of online distance education, Twidale and Ruhleder [75] pointed out several design

implications for conducting over-the-shoulder learning, such as maintaining context and enabling multi-application switching, which may be informative for future work on open-ended studios offered online at scale in a university setting. Studies around remote collaborative learning in engineering labs [52, 64] or museums [3] may also offer inspiration to actualize over-the-shoulder peer learning in online studios via shared context digitally.

In our interviews, instructors considered increased diversity in peers' work and ideas as an opportunity to enhance peer learning via work-sharing and discussion. Some art-focused MOOC platforms allow students to showcase their work through an e-portfolio, e.g., Kadenze [65], to allow students to learn from each other's work. Given a large number of deliverables from peers with simple ordering options based on popularity or recentness, students may experience information-overload and have trouble locating those that help their learning the most [39]. Increased visibility of peers' work may also entice visual plagiarism [77]. Future studies that create or improve platforms for work-sharing among peers should keep these issues in mind.

6.3 Adapt to Open-ended Assessment

Based on our surveys, for open-ended courses, close-ended assignments and standardized tests were not the best ways to learn or to evaluate in many cases (Section 4.1.2, Section 4.2.2). Open-ended assignments and exams were more prevalent in open-ended courses. Instructors reported feeling overwhelmed with their workload–maintaining the same number of open-ended assignments and the same degree of openness on assignments.

Today, many MOOCs typically rely on quizzes and assignments with close-ended questions (e.g., multiple-choice) or simple open-ended questions (e.g., short answers) as assessment, due to their easier operationalization and grading [14]. Kadenze, an art-focused MOOC platform [61], includes open-ended artifact creation assignments in various forms, such as music recordings, drawings, and graphic design. However, instructors only grade assignments of those students who pay for the course; others are evaluated through peer assessment. To the best of our knowledge, there is no existing work on the auto-grading of artifact creation assignments.

For open-ended argument-oriented assignments that involve posting in a discussion forum and responding to peers, based on the instructor interviews, the number of posts was a common metric for grading in a large online course. This method, however, may lead to low student engagement, according to the instructors (Section 5.1). Automated Essay Scoring is sometimes used in grading open-ended argument-oriented assignments [4]. This evaluation approach focuses on evaluating writing quality rather than assessing the quality of opinions or arguments. As a result, it is limiting in our context because the content in open-ended courses requires introspection and criticism, and students' proficiency level in the language of instruction could vary in a diverse student body.

Thus, we need more research to support the efficient grading of more types of open-ended assignments, such as artifact creation and argument-oriented writing. The grading criteria should respect the openness of these assignments and focus on qualities such as the level of introspection and creativity. Again, university policy-makers should understand the time required for instructors to grade open-ended assignments and provide additional resources to support them.

6.4 Scale Individual Open-ended Feedback

Feedback plays an indispensable role in learning [36]. Due to the frequent use of open-ended assignments in open-ended courses and the need for specific and immediate feedback, according to our surveys and interviews (Section 4.1.3, Section 4.2.3, Section 5.3), instructors faced a time-quality trade-off on providing individual feedback in a massive course.

Nowadays, MOOC platforms mainly utilize peer review to provide feedback to complex, open-ended work [70]. While there is a body of work that targets the improvement of peer feedback in

creative work [11, 28, 41, 47], our survey and interview results suggest that peer review should not be the sole feedback channel because of the inconsistent quality of the feedback. Moreover, the platforms alone do not dictate the quality of peer review. Guidelines in the form of rubrics and norm-setting by the instructors are critical [47]. Seeing what an instructor values as they participate in the peer review process becomes a learning experience and mitigates ambiguity for student peers.

As far as we know, there is no existing method that provides automatic customized feedback to open-ended assignments, beyond assessing writing-quality. The most a computer algorithm can do currently on a complicated open-ended assignment is assigning a score and a ranking [30]. Open-ended artifacts and long content-based natural language responses have such a high degree of complexity that the feasibility of automating feedback is an open question. Due to the complexity of open-ended assignments, human-in-the-loop algorithms might be a more feasible solution in the near term. As an example, Brooks et al. [10] assisted instructors in providing customized feedback to short-answer questions on a large scale using a clustering approach. Yen et al. [84] developed a visualization tool to help students better organize and interpret feedback from multiple crowd-sourced feedback providers.

6.5 Establish Trust for Self-expression

Based on our interviews and surveys, learning tasks and assignments in open-ended courses incorporated more personal opinions, experience, and emotion than non-open-ended courses due to their value-laden nature. As a result, instructors of open-ended courses highly valued the establishment of trust among students (Section 4.1.3), and students preferred an open, unbiased environment for comfortable self-expression in open-ended courses (Section 4.2.2). Ossiannilsson et al. (2015) [59] confirmed trust as a critical element for forming a quality culture on MOOCs.

In the interviews, instructors of large open-ended courses reported challenges in establishing trust and maintaining a welcoming environment due to increased anonymity and evaluation apprehension. Instructors of online courses also had a hard time fostering such an environment due to difficulty in breaking formalities, creating a sense of community, and having emotional exchanges. Open-ended courses offered online and at scale, having a large student body and being online simultaneously, face both of these challenges.

Current MOOC platforms have few tools or measures to address issues of trust and inclusion in open-ended courses. Often, MOOC platforms have a Code of Conduct to discourage disrespectful or offensive language. Most of the time, they rely on human moderators to enforce the rules in the discussion forums, which is time-consuming and costly. Some of them include self-introductions at the beginning of the course to increase social presence.

While researchers have shown that trust is one of the key enablers of knowledge sharing in online communities [79], less work exists on practical ways to establish trust and create an inclusive environment for large scale online courses. There is a rich body of work that studies the practice of moderation and norm-setting in large-scale online communities, such as social media platforms and interest-based online forums, [33, 63, 45], which aims to create a welcoming environment, but little is in an educational context. The goal of moderation could be different between an online course versus a non-educational-centric community. The only available work related to moderation in education focuses mostly on in-person or small online courses [74, 44]. Moderation in in-person or smaller online courses may not apply to large-scale online courses because of the differences in scale and affordances. Nevertheless, future research may draw inspiration from theories and findings in the moderation area and apply them to establishing trust in open-ended courses offered online at scale.

In summary, much more research is needed to address *how* to create a welcoming environment and set up norms that encourage students to self-express in front of a broad audience, potentially with a drastically different background. We suggest developing proactive measures that guide healthy debates and critiques and create more opportunities to establish trust and a sense of community.

6.6 Personalize Instruction and Harness the Benefits of Diversity

According to our surveys and interviews, incorporating students' backgrounds into personalized instruction was a critical trait of open-ended courses (Section 4.1.1, Section 5.5). In addition to individualized feedback, as discussed above, other parts of the instructional process also needed personalization.

At present, major MOOC platforms do not allow for personalized instruction. Nevertheless, prior work has tried to customize learning paths based on student schedule, activity amount, mastery level, prior knowledge, learning style, and learning goal [71]. One significant difference between open-ended and non-open-ended courses is whether knowledge can be modularized [53]. In non-open-ended courses, content can be compartmentalized into knowledge components that are clearly defined. In open-ended courses, students engage with a body of knowledge or practice that is fluid. As a result, certain personalization approaches in large-scale online non-open-ended courses may not apply to the open-ended courses. For example, while the system can adjust the learning path based on students' mastery level on learning the knowledge components in non-open-ended courses, it is harder for open-ended courses to define if a student has mastered a knowledge component. Future research should consider other types of personalization that better suit open-ended subjects, such as changing the manner of explaining value-laden topics that fit the student's background, and adjusting the evaluation standard as needed.

While personalization is a significant challenge in a massive class with diverse students, such diversity could be an asset for open-ended courses at scale, as suggested in our interviews (Section 5.1) and in prior work [46]. Early MOOCs, cMOOCs, were based on the learning theory of connectivism, and they emphasized the potential of knowledge co-construction and sharing amongst students [24]. We suggest that the application of connectivism learning theory can potentially harness the benefits of diversity in large-scale online open-ended courses. Recent work on learner-sourcing [80, 78] is also offering novel learning and assessment opportunities that benefit from a diverse student population. For future research, there is a significant opportunity to develop tools and systems that can create, organize, and utilize student-created content from different parts of a large-scale online course, such as assignments and discussion forums.

6.7 Researcher's Stance on Learning-at-scale

Researchers of this study advocate for positive learning outcomes and have kept a neutral position towards learning-at-scale throughout the study. While we recognize the benefits of enhancing the accessibility to higher education, we warn against the phenomenon in which universities and colleges over-scale their offerings to an unmanageable size. Universities should carefully consider the above challenges and concerns raised by instructors and students and allocate sufficient resources, design relevant policies, and leverage appropriate educational technologies to deliver quality learning outcomes at scale.

7 LIMITATIONS

In this section, we identify limitations in our study from three perspectives, quality of self-reporting, sample selection, and the generalizability of the study.

7.1 Quality of self-reporting

Our study relied heavily on self-reported data. We discuss two limitations of self-reported data below.

Self-reporting bias: Self-reporting bias is a widely-known pitfall of surveys and interviews [2], including recall bias and recency bias. It is encouraging that four of the five pedagogical elements we identified were supported by findings from both the instructor survey and the student survey. Nevertheless, future studies may use approaches not based on self-reporting, such as observational study or randomized controlled trials, to verify if our findings based on self-reporting are biased. For example, researchers may observe and compare the learning outcome of two classes, one encouraging self-expression while the other does not, in a semi-controlled environment, to confirm if self-expression is indeed a critical element for effective teaching and learning in open-ended courses.

Potential misalignment between perceived needs and evidence: Previous work by Deslauriers et al. [23] demonstrated the possibility of misalignment between students' perceived effectiveness of a type of pedagogy and its actual effect based on learning outcome. Similarly, in our study, pedagogy elements that students or instructors preferred in their responses bear the risk of not being effective in practice. However, it is encouraging that four of the five pedagogical elements we identified were supported by findings from both the instructor survey and the student survey. Our study is exploratory and aims to identify future research opportunities.

7.2 Sample Selection

As a mixed-methods study with interviews and two surveys, the sampling process was complicated. Despite our best efforts to collect representative samples, below, we discuss two limitations related to low survey response and a small interview sample.

Low survey response rates: In our instructor and student surveys, a response rate of 5.04% and 2.62% respectively were relatively low, which is a common phenomenon. To identify selection bias, we compared the samples' demographics distribution with that of the university's overall faculty population and student population. The most apparent difference was a lower female percentage in our instructor sample than the faculty population (36% vs. 49%). Other features such as academic title, department align with the overall faculty population. The demographics between the instructors of open-ended courses and the instructors of non-open-ended courses were similar, as shown in Section 3. As for a non-response bias, it was not as prevalent in the instructor survey since we performed a between-subject comparison, where both groups were subjected to a similar non-response bias. We acknowledge a potential non-response bias in the student sample. One possible scenario is that students who responded to the survey had stronger opinions about open-ended courses than those who did not. As a result, their preferences may not represent the entire student population. Future work can validate the findings with samples that are closer to random.

Limited interview sample: In our interview, we focused on recruiting instructors who had taught large or online open-ended courses based on a theoretical sampling approach [51]. Since a substantial proportion of open-ended courses in the university are small and in-person, there was a limited pool of instructors from which we could sample. Despite our best effort, the final interview sample is relatively small. Nevertheless, prior studies have shown that interview code saturation can be achieved with 12 or as few as six interviews, given that the interview is relatively structured, the interviewees have independent experiences, and are relatively heterogeneous [34]. Our interview satisfies these assumptions (please see Section 3.3 for

details). In addition, our interview sample does not include instructors from all departments that offered mostly open-ended courses; there are forty-five departments of this kind. As a result, the challenges and opportunities we identified may not be exhaustive and may not generalize across departments. Future studies should include a larger number of instructors who teach open-ended courses from a more extensive range of departments to confirm and expand our findings.

7.3 Generalizability of Findings

Our study could not cover all possible types of courses and universities, and we recognize the limitations on the generalizability of our findings.

University settings: All instructors and students in our survey and interview samples reported their experiences in a university setting. Therefore, our findings may not generalize to education at scale for open-ended courses in non-university-settings, such as MOOCs. MOOCs have a lower entry barrier, and students taking these courses often have different learning goals and a lower level of engagement than those in a university setting. Instructors may face additional challenges in teaching these courses at scale. Nevertheless, our findings still offer implications for offering open-ended courses at scale in non-university settings, but require further confirmation.

Single university: Our sampled participants for surveys and interviews are from a single, large public research (R1) university. We did not include instructors and students from other universities, including liberal arts colleges, teaching colleges, and private universities. Thus, our findings may not generalize across different university types as institutions may differ in their pedagogical goals, philosophies, and instructional resources. Future work should expand our investigation to represent other university types.

Class size: The size of the large-scale courses discussed in our interviews ranged between 30 to 300 students. While we expect many of our findings to persist at a larger scale, there may be challenges at a larger scale that our interviewees did not capture.

8 CONCLUSION

Our work is the first to systematically compare differences in teaching open-ended and non-open-ended courses in a university setting, and identify concerns that instructors and students have about scaling open-ended classes online in the Arts, Humanities and Social Sciences. Our mixed-methods study, incorporating surveys of instructors and students, and semi-structured instructor interviews, identifies five critical pedagogy elements that instructors valued as necessary in open-ended courses, but not as necessary or applicable for non-open-ended courses. An overarching theme for the five elements was the need to support students' self-expression in open-ended courses. We further uncover the open challenges and opportunities for incorporating these pedagogical elements that support effective teaching of open-ended courses at scale and online in a university setting. Finally, we suggest six future research directions based on the open challenges and opportunities: (1) facilitate in-depth conversations, (2) create a studio-friendly environment, (3) adapt to open-ended assessment, (4) scale individual open-ended feedback, (5) establish trust for self-expression, and (6) personalize instruction and harness the benefits of student diversity.

We conducted this study during April–December 2019, before the COVID-19 pandemic. The outbreak resulted in a forced mass migration to online pedagogy [85]. Experience with online pedagogy will likely increase after Spring 2020. We hope this work provides guidance to instructors as they migrate their university-level open-ended courses to online environments, and to those who

hope to increase accessibility to higher-education via offering high-quality open-ended courses in a learning-at-scale environment in the future.

ACKNOWLEDGMENTS

We would like to thank all the instructors and students who voluntarily participated in our study, Meng Huang for his help in streamlining the survey analysis, and members of the Social Spaces Group and Crowd Dynamics Lab for their feedback.

REFERENCES

[1] Carlos Alario-Hoyos, Mar Pérez-Sanagustín, Carlos Delgado-Kloos, Mario Muñoz-Organero, Antonio Rodríguez-de-las-Heras, et al. 2013. Analysing the impact of built-in and external social tools in a mooc on educational technologies. In *European Conference on Technology Enhanced Learning*. Springer, 5–18.

[2] Alaa Althubaiti. 2016. Information bias in health research: definition, pitfalls, and adjustment methods. *Journal of multidisciplinary healthcare*, 9, 211.

[3] Ernesto Arroyo, Valeria Righi, Roger Tarrago, Patricia Santos, Davinia Hernández-Leo, and Josep Blat. 2011. Remote collaborative multi-user informal learning experiences: design and evaluation. In *European Conference on Technology Enhanced Learning*. Springer, 43–56.

[4] Stephen P Balfour. 2013. Assessing writing in moocs: automated essay scoring and calibrated peer review™. *Research & Practice in Assessment*, 8, 40–48.

[5] William Bianchi. 2002. The wisconsin school of the air: success story with implications. *Journal of Educational Technology & Society*, 5, 1, 141–147.

[6] Patrick Blessinger and John M Carfora. 2014. *Inquiry-based learning for the arts, humanities and social sciences: A conceptual and practical resource for educators*. Emerald Group Publishing.

[7] Benjamin S Bloom et al. 1956. Taxonomy of educational objectives. vol. 1: cognitive domain. *New York: McKay*, 20–24.

[8] College Board. 2020. Quick guide: types of college courses. https://bigfuture.collegeboard.org/find-colleges/academic-life/quick-guide-types-of-college-courses. (2020).

[9] Sorana BolboacÄ, JÃ Lorentz, et al. 2007. Computer-based testing on physical chemistry topic: a case study. *International Journal of Education and Development using ICT*, 3, 1.

[10] Michael Brooks, Sumit Basu, Charles Jacobs, and Lucy Vanderwende. 2014. Divide and correct: using clusters to grade short answers at scale. In *Proceedings of the first ACM conference on Learning@ scale conference*. ACM, 89–98.

[11] Julia Cambre, Scott Klemmer, and Chinmay Kulkarni. 2018. Juxtapeer: comparative peer review yields higher quality feedback and promotes deeper reflection. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, 294.

[12] David Carless. 2009. Trust, distrust and their impact on assessment reform. *Assessment & Evaluation in Higher Education*, 34, 1, 79–89.

[13] Muthu Kumar Chandrasekaran, Min-Yen Kan, Bernard CY Tan, and Kiruthika Ragupathi. 2015. Learning instructor intervention from mooc forums: early results and issues. *arXiv preprint arXiv:1504.07206*.

[14] Jyoti Chauhan and Anita Goel. 2016. An analysis of quiz in mooc. In *2016 Ninth International Conference on Contemporary Computing (IC3)*. IEEE, 1–6.

[15] Mary M Christopher, Julie A Thomas, and Mary K Tallent-Runnels. 2004. Raising the bar: encouraging high level thinking in online discussion forums. *Roeper Review*, 26, 3, 166–171.

[16] Derrick Coetzee, Armando Fox, Marti A Hearst, and Bjoern Hartmann. 2014. Chatrooms in moocs: all talk and no action. In *Proceedings of the first ACM conference on Learning@ scale conference*. ACM, 127–136.

[17] R Cook. 2000. Technology trends: teleconferencing, distance education, and future technologies. (2000).

[18] Claude Cookman. 2009. Using jitt to foster active learning in a humanities course. *Just-in-Time Teaching*, 163–178.

[19] Yi Cui and Alyssa Friend Wise. 2015. Identifying content-related threads in mooc discussion forums. In *L@ S*, 299–303.

[20] Thanasis Daradoumis, Roxana Bassi, Fatos Xhafa, and Santi Caballé. 2013. A review on massive e-learning (mooc) design, delivery and assessment. In *2013 eighth international conference on P2P, parallel, grid, cloud and internet computing*. IEEE, 208–213.

[21] Brian Dear. 2017. *The friendly orange glow: The untold story of the PLATO System and the Dawn of Cyberculture*. Pantheon.

[22] Vanessa Paz Dennen*. 2005. From message posting to learning dialogues: factors affecting learner participation in asynchronous discussion. *Distance Education*, 26, 1, 127–148.

[23] Louis Deslauriers, Logan S McCarty, Kelly Miller, Kristina Callaghan, and Greg Kestin. 2019. Measuring actual learning versus feeling of learning in response to being actively engaged in the classroom. *Proceedings of the National Academy of Sciences*, 116, 39, 19251–19257.

[24] Stephen Downes. 2008. Places to go: connectivism & connective knowledge. *Innovate: Journal of Online Education*, 5, 1, 6.

[25] Suzannah Evans and Karen McIntyre. 2016. Moocs in the humanities: can they reach underprivileged students? *Convergence*, 22, 3, 313–323.

[26] Suzannah Evans and Jessica Gall Myrick. 2015. How mooc instructors view the pedagogy and purposes of massive open online courses. *Distance Education*, 36, 3, 295–311.

[27] Helene Fournier and Rita Kop. 2015. Mooc learning experience design: issues and challenges. *International Journal on E-Learning*, 14, 3, 289–304.

[28] C Ailie Fraser, Tricia J Ngoon, Ariel S Weingarten, Mira Dontcheva, and Scott Klemmer. 2017. Critiquekit: a mixed-initiative, real-time interface for improving feedback. In *Adjunct Publication of the 30th Annual ACM Symposium on User Interface Software and Technology*. ACM, 7–9.

[29] Nicola K Gale, Gemma Heath, Elaine Cameron, Sabina Rashid, and Sabi Redwood. 2013. Using the framework method for the analysis of qualitative data in multi-disciplinary health research. *BMC medical research methodology*, 13, 1, 117.

[30] Chase Geigle, ChengXiang Zhai, and Duncan C Ferguson. 2016. An exploration of automated grading of complex assignments. In *Proceedings of the Third (2016) ACM Conference on Learning@ Scale*. ACM, 351–360.

[31] Graham R Gibbs. 2007. Thematic coding and categorizing. *Analyzing qualitative data. London: Sage*, 38–56.

[32] Joshua Goodman, Julia Melkers, and Amanda Pallais. 2019. Can online delivery increase access to education? *Journal of Labor Economics*, 37, 1, 1–34.

[33] James Grimmelmann. 2015. The virtues of moderation. *Yale JL & Tech.*, 17, 42.

[34] Greg Guest, Arwen Bunce, and Laura Johnson. 2006. How many interviews are enough? an experiment with data saturation and variability. *Field methods*, 18, 1, 59–82.

[35] Kathleen Harting and Margaret J Erthal. 2005. History of distance learning. *Information technology, learning, and performance journal*, 23, 1, 35.

[36] John Hattie and Helen Timperley. 2007. The power of feedback. *Review of educational research*, 77, 1, 81–112.

[37] Donald E Heller. 2001. *The states and public higher education policy: Affordability, access, and accountability*. JHU Press.

[38] Khe Foon Hew and Wing Sum Cheung. 2014. Students' and instructors' use of massive open online courses (moocs): motivations and challenges. *Educational research review*, 12, 45–58.

[39] Starr R Hiltz and Murray Turoff. 1985. Structuring computer-mediated communication systems to avoid information overload. *Communications of the ACM*, 28, 7, 680–689.

[40] Malinka Ivanova. 2016. Technology landscape in moocs platforms. In *2016 19th International Symposium on Electrical Apparatus and Technologies (SIELA)*. IEEE, 1–4.

[41] Hyeonsu B Kang, Gabriel Amoako, Neil Sengupta, and Steven P Dow. 2018. Paragon: an online gallery for enhancing design feedback with visual examples. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, 606.

[42] Desmond Keegan. 2013. *Foundations of distance education*. Routledge.

[43] Hanan Khalil and Martin Ebner. 2013. "how satisfied are you with your mooc?"-a research study on interaction in huge online courses. In *EdMedia Innovate Learning*. Association for the Advancement of Computing in Education (AACE), 830–839.

[44] Andrea Kienle and Carsten Ritterskamp. 2007. Facilitating asynchronous discussions in learning communities: the impact of moderation strategies. *Behaviour & Information Technology*, 26, 1, 73–80.

[45] Sara Kiesler, Robert Kraut, Paul Resnick, and Aniket Kittur. 2012. Regulating behavior in online communities. *Building successful online communities: Evidence-based social design*, 125–178.

[46] Chinmay Kulkarni, Julia Cambre, Yasmine Kotturi, Michael S Bernstein, and Scott R Klemmer. 2015. Talkabout: making distance matter with small groups in massive classes. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*. ACM, 1116–1128.

[47] Chinmay E Kulkarni, Michael S Bernstein, and Scott R Klemmer. 2015. Peerstudio: rapid peer feedback emphasizes revision and improves performance. In *Proceedings of the second (2015) ACM conference on learning@ scale*. ACM, 75–84.

[48] Fabrizio Lamberti, Andrea Sanna, Gianluca Paravati, and Gilles Carlevaris. 2014. Automatic grading of 3d computer animation laboratory assignments. *IEEE Transactions on Learning Technologies*, 7, 3, 280–290.

[49] N Levenburg. 2000. A brief history of curriculum reform in the united states. *Distance Learning Dynamics*.

[50] Ou Lydia Liu, Chris Brew, John Blackmore, Libby Gerard, Jacquie Madhok, and Marcia C Linn. 2014. Automated scoring of constructed-response science items: prospects and obstacles. *Educational Measurement: Issues and Practice*, 33, 2, 19–28.
[51] Karen D Locke. 2000. *Grounded theory in management research*. Sage.
[52] Jan Machotka, Zorica Nedic, and Özdemir Göl. 2008. *Collaborative learning in the remote laboratory NetLab*. PhD thesis. International Institute of Informatics and Systemics.
[53] Philip W Martin. 2003. Key aspects of teaching and learning in arts, humanities and social sciences. *A handbook for teaching and learning in higher education*, 300.
[54] Michael J McInerney and L Dee Fink. 2003. Team-based learning enhances long-term retention and critical thinking in an undergraduate microbial physiology course. *Microbiology Education*, 4, 3.
[55] John A Michael. 1980. Studio art experience: the heart of art education. *Art Education*, 33, 2, 15–19.
[56] Michael Mohler, Razvan Bunescu, and Rada Mihalcea. 2011. Learning to grade short answer questions using semantic similarity measures and dependency graph alignments. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies-Volume 1*. Association for Computational Linguistics, 752–762.
[57] Michael G Moore. 1989. Three types of interaction. (1989).
[58] National Association of Schools of Art and Design. 2018. Nasad handbook 2018-19. https://nasad.arts-accredit.org/accreditation/standards-guidelines/handbook/. (Dec. 2018).
[59] Ebba Ossiannilsson, Fahriye Altinay, and Zehra Altinay. 2015. Analysis of moocs practices from the perspective of learner experiences and quality culture. *Educational Media International*, 52, 4, 272–283.
[60] Tina Overton. 2003. Key aspects of teaching and learning in experimental sciences and engineering. *A handbook for teaching and learning in higher education: Enhancing academic practice*, 2, 255–277.
[61] Wanqi Peng. 2016. How can mooc providers create an interactive learning experience in the arts. (2016).
[62] Justin Reich and José A Ruipérez-Valiente. 2019. The mooc pivot. *Science*, 363, 6423, 130–131.
[63] Yuqing Ren, Robert Kraut, Sara Kiesler, and Paul Resnick. 2010. Regulating behavior in online communities.
[64] Frederico Menine Schaf, Dieter Müller, F Wilhelm Bruns, Carlos Eduardo Pereira, and H-H Erbe. 2009. Collaborative learning and engineering workspaces. *Annual Reviews in Control*, 33, 2, 246–252.
[65] D Shah. 2018. Kadenze programs: a modern day credential for artists and creatives. (2018).
[66] Dhawal Shah. 2019. By the numbers: moocs in 2019. https://www.classcentral.com/report/mooc-stats-2019/. (Dec. 2019).
[67] Peter Shea, Alexandra Pickett, and Chun Sau Li. 2005. Increasing access to higher education: a study of the diffusion of online teaching among 913 college faculty. *The International review of research in open and distributed learning*, 6, 2.
[68] Ali Sher. 2009. Assessing the relationship of student-instructor and student-student interaction to student learning and satisfaction in web-based online learning environment. *Journal of Interactive Online Learning*, 8, 2.
[69] Scott Simkins and Mark Maier. 2010. *Just-in-time teaching: Across the disciplines, across the academy*. Stylus Publishing, LLC.
[70] Hoi K Suen. 2014. Peer assessment for massive open online courses (moocs). *International Review of Research in Open and Distributed Learning*, 15, 3, 312–327.
[71] Ayse Saliha Sunar, Nor Aniza Abdullah, Su White, and Hugh C Davis. 2015. Personalisation of moocs: the state of the art.
[72] Karen Swan, Scott Day, Leonard Bogle, and Traci van Prooyen. 2015. Amp 1: a tool for characterizing the pedagogical approaches of moocs. In *MOOCs and open education around the world*. Routledge, 105–118.
[73] Michael Sweet and Larry K Michaelsen. 2012. *Team-based learning in the social sciences and humanities: Group work that works to generate critical thinking and engagement*. Stylus Publishing, LLC.
[74] Joan Thormann and Patricia Fidalgo. 2014. Guidelines for online course moderation and community building from a student's perspective. *Journal of Online Learning & Teaching*, 10, 3, 374–388.
[75] Michael B Twidale and Karen Ruhleder. 2004. Over-the-shoulder-learning in a distance education environment. *Learning, culture, and community in online education: research and practice, Caroline Haythornthwaite & Michelle M. Kazmer, eds. New York, NY: Lang*, 177–194.
[76] Arizona State University. 2019. Defining course components. https://provost.asu.edu/sites/default/files/page/1585/defining-course-components_6-17-19.pdf. (June 2019).
[77] Beth Walker. 2009. New twists on an old problem: preventing plagiarism and enforcing academic integrity in an art and design school. *Art Documentation: Journal of the Art Libraries Society of North America*, 28, 1, 48–51.
[78] Xu Wang, Srinivasa Teja Talluri, Carolyn Rose, and Kenneth Koedinger. 2019. Upgrade: sourcing student open-ended solutions to create scalable learning opportunities. In *Proceedings of the Sixth (2019) ACM Conference on Learning@Scale*, 1–10.

[79] Etienne Wenger, Nancy White, and John D Smith. 2009. *Digital habitats: Stewarding technology for communities*. CPsquare.
[80] Joseph Jay Williams, Juho Kim, Anna Rafferty, Samuel Maldonado, Krzysztof Z Gajos, Walter S Lasecki, and Neil Heffernan. 2016. Axis: generating explanations at scale with learnersourcing and machine learning. In *Proceedings of the Third (2016) ACM Conference on Learning@ Scale*, 379–388.
[81] Pete Williams, David Nicholas, and Barrie Gunter. 2005. E-learning: what the literature tells us about distance education. In *Aslib Proceedings*. Emerald Group Publishing Limited.
[82] Zhen Xiong, Lin Zhi, and Jie Jiang. 2019. Research on art education digital platform based on big data. In *2019 IEEE 4th International Conference on Big Data Analytics (ICBDA)*. IEEE, 208–211.
[83] Diyi Yang, Mario Piergallini, Iris Howley, and Carolyn Rose. 2014. Forum thread recommendation for massive open online courses. In *Educational Data Mining 2014*. Citeseer.
[84] Yu-Chun Grace Yen, Joy O Kim, and Brian P Bailey. 2020. Decipher: an interactive visualization tool for interpreting unstructured design feedback from multiple providers. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1–13.
[85] Jeff Young. 2020. Sustaining higher education in the coronavirus crisis. *Edsurge*. Accessed 6/1/2020 at: https://www.edsurge.com/research/guides/sustaining-higher-education-in-the-coronavirus-crisis.
[86] Yuan Zhenming, Zhang Liang, and Zhan Guohua. 2003. A novel web-based online examination system for computer science education. In *33rd ASEE/IEEE Frontiers in Education Conference*, 5–8.
[87] Craig B. Zilles, Matthew West, Geoffrey L. Herman, and Timothy Bretl. 2019. Every university should have a computer-based testing facility. In *CSEDU*.

A METHODS

A.1 Definition of course components in the instructor survey

In the instructor survey, we asked instructors about the course components they had in their courses and how they conducted each component. For each component, we asked about the amount of instructor-student interaction and peer interaction included in the component and the interaction channels used. We included three course components based on their definitions in universities and colleges [76, 8].

- **A lecture-based session** primarily contains an oral presentation by an instructor but may involve various active learning activities such as think-pair-share, polling.
- **A discussion or a seminar session**, compared to a lecture-based session, focuses more on two-way exchanges of ideas via conversations.
- **A lab session or a studio session**, compared to a discussion or a seminar session, emphasizes the development of an artifact or completion of a specific task.

While a lecture session often accompanies a discussion session, but a seminar is often standalone, a discussion and a seminar share the characteristics that both have smaller session sizes and are designed to involve more interaction, which distinguishes them from a lecture session. A lab session is often in STEM subjects while a studio is often in the Arts. Nevertheless, we grouped them because they are both defined as hands-on workshops, just for different subjects.

A.2 Sample Factors Comparison

As mentioned in the Methods section (Section 3.1), we compared the pedagogical choices made by instructors between open-ended and non-open-ended courses based on the responses of the instructor survey. We tried our best to match the samples of open-ended and non-open-ended courses along the dimensions that may influence pedagogy decisions between the two comparison groups. We checked the percentage of online courses, the distribution of different class sizes, the instructor's academic rank, gender, class difficulty level, and the highest learning goal categorized by Bloom's taxonomy [7].

Table 6. Sample Statistics Comparison between Open-ended vs. Non-open-ended Courses

	Open-ended	**Non-open-ended**
# of Responses	61	60
# of Departments	45	40
% Online	34.43%	31.67%
Distribution of Class Sizes	34.42% small, 29.51% medium, 36.07% large	40% small, 20% medium, 40% large
Distribution of Instructor's Academic Rank	About 39.34% Professor, 22.95% Associate Prof., 21.31% Assistant Prof., 16.4% Others	About 50% Professor, 18.33% Associate Prof., 21.67% Assistant Prof., 10% Others
Distribution of Instructor's Gender	59.02% Man, 41.67% Woman	65% Man, 33.33% Woman, 1.67% Others
Distribution of Class Level	22.95% Undergrad Intro, 14.75% Undergrad Intermediate, 29.51% Undergrad Advanced, 32.79% Graduate-level	25% Undergrad Intro, 8.33% Undergrad Intermediate, 33.33% Undergrad Advanced, 33.33% Graduate-level
Distribution of the Highest Learning Goal	44.26% Applying, 31.15% Understanding, 21.31% Analyzing, Evaluating, and Creating	60% Applying, 16.67% Understanding, 19.67% Analyzing, Evaluating, and Creating

In Table 6, we can see that the open-ended and non-open-ended course samples are very similar in all of these factors, suggesting that there is no significant difference between the open-ended courses and non-open-ended courses samples to the best of our knowledge.

Received June 2020; accepted July 2020