



Interest in Educational Data and Barriers to Data Use Among Massive Open Online Course Instructors

Maya Usher¹ · Arnon Hershkovitz¹

Accepted: 6 July 2022 / Published online: 6 August 2022
© The Author(s), under exclusive licence to Springer Nature B.V. 2022

Abstract

Higher education instructors constantly rely on educational data to assess and evaluate the behavior of their students and to make informed decisions such as which content to focus on and how to best engage the students with it. Massive open online course (MOOC) platforms may assist in the data-driven instructional process, as they enable access to a wide range of educational data that is gathered automatically and continuously. Successful implementation of a data-driven instruction initiative depends highly on the support and acceptance of the instructors. Yet, our understanding of instructors' perspectives regarding the process of data-driven instruction, especially with reference to MOOC teaching, is still limited. Hence, this study was set to characterize MOOC instructors' interest in educational data and their perceived barriers to data use for decision-making. Taking a qualitative approach, data were collected via semi-structured interviews with higher education MOOC instructors from four public universities in Israel. Findings indicated that the instructors showed great interest mostly in data about social interactions between learners and about problems with the MOOC educational resources. The main reported barriers for using educational data for decision-making were lack of customized data, real-time access, data literacy, and institutional support. The results highlight the need to provide MOOC instructors with professional development opportunities for the proper use of educational data for skilled decision-making.

Keywords Data-driven decision-making · Higher education · Instructor perceptions · Massive open online courses (MOOCs) · Online teaching

Introduction

Higher education instructors are constantly using educational data to understand and evaluate the behavior of their class and individual students (Leitner et al., 2017; Picciano, 2012). Collecting and analyzing educational data with the purpose to guide and support educational decisions is known as the process of data-driven decision-making (Prinsloo & Slade, 2014). Such decisions may be related to the topics the instructors should put more emphasis on, potential ways to engage the students during classes, and evaluation of learning outcomes (Maisarah et al., 2021).

In face-to-face courses, instructors are accustomed to observing educational data, and responding to them, relying on both verbal and non-verbal cues (Herodotou et al., 2019; Vanlommel et al., 2017). However, with the transition to online remote teaching, as occurred during the COVID-19 outbreak, the way instructors and students interact has significantly changed, and as a result, so has the range of data to which instructors are exposed (Maisarah et al., 2021; Usher, Hershkovitz & Forkosh-Baruch, 2021b). While teaching online, instructors experience indirect interaction with students; hence, they are less exposed to non-verbal data that is continuously available in the physical classroom (Herodotou et al., 2019; Usher et al., 2021b).

This is especially true for massive open online courses (MOOCs). MOOCs are web-based learning environments, designed to provide free and accessible high-quality education to the masses (Margaryan et al., 2015). Due to the huge amount of users, and to the mostly asynchronously instructor-student interaction, MOOCs pose some challenges for open and direct communication and for personalized

✉ Maya Usher
oshermaya@gmail.com
Arnon Hershkovitz
arnonhe@tauex.tau.ac.il

¹ Department of Mathematics, Science, and Technology Education, School of Education, Tel Aviv University, 6997801 Tel Aviv, Israel

instruction (Ruipérez-Valiente et al., 2017; Wang & Woo, 2007; Usher et al., 2021a). This situation compromises the data-driven instructional process since the instructors rely on only a thin layer of easily observed student course activity, mainly grades (Alexandron et al., 2019; Gašević et al., 2016).

Still, online learning systems may assist in the data-driven instructional process as they usually keep digital traces left by learners while engaging with the course material and assignments (Vigentini et al., 2017). The digital traces left by MOOC learners enable access to a wide range of educational data that is gathered automatically and continuously (Er et al., 2019; Vigentini et al., 2017). Collecting and analyzing such data for purposes of understanding and optimizing learning and the environments in which it occurs helped to inform the rise of the field of learning analytics (LA) (Larrabee Sønderlund et al., 2019; Siemens, 2013).

Yet, mere access to educational data is not enough. Successful implementation of data-driven instruction requires mastery of many twenty-first-century high-level thinking skills, and analytical thinking in particular (Green et al., 2016; OECD, 2018; Raffaghelli & Stewart, 2020). With reference to higher education, analytical thinking refers to the ability to extract key information from large data sets about students' activity and make informed decisions based on the collected data (Yulina et al., 2019). Yet, recent studies have reported on the many challenges instructors are facing while implementing data-driven instruction in their courses, one of the most common of which was lack of core data literacies (Hilliger et al., 2020; Klein et al., 2019; Shibani et al., 2020).

Moreover, successful implementation of data-driven instruction depends highly on the support and acceptance of the instructors (Leitner et al., 2017; Vigentini et al., 2017). This is problematic since most designers of LA tools tend to focus on technical considerations, rather than on stakeholders' desires and actual needs (Hilliger et al., 2020; Holstein et al., 2019). Therefore, it is not surprising that according to instructors' reports, another common challenge they are facing while implementing LA tools is that the data provided to them is not aligned to their needs and pedagogical perspectives (Klein et al., 2019; Shibani et al., 2020). This observation has led recent studies to explore new methods and strategies for co-designing LA tools with critical stakeholders, such as the teachers (An et al., 2020; Holstein et al., 2019).

Recently, studies have begun examining higher education instructors' perspectives regarding the use of educational data, concluding that the major barriers for LA adoption include lack of core data literacies, lack of timely information, and lack of personalization data (Hilliger et al., 2020; Klein et al., 2019; Shibani et al., 2020). Most studies in this field have been undertaken in the context of face-to-face courses that use online learning management systems. There

is a lack of studies that specifically address the perspectives of instructors who teach in MOOCs regarding the data-driven instructional process (Er et al., 2019). The current study aims at bridging this gap.

The goal of the current study was to characterize MOOC instructors' interest in educational data and their perceived barriers to data use. To meet this goal, the following research questions were explored:

1. What types of educational data are of interest to MOOC instructors?
2. What are MOOC instructors' perceived barriers to data use for decision-making?

Literature Review

Data-Driven Instruction in Higher Education

Higher education instructors constantly use educational data (such as students' grades and participation patterns) to assess and evaluate the behavior of their class and individual students (Leitner et al., 2017; Picciano, 2012). Based on educational data, instructors can create meaningful insights for actionable decisions, such as which content to focus on, how to best engage the students with it, and how to evaluate students' learning outcomes (Maisarah et al., 2021). This process of making meaningful decisions based on educational data is known as data-driven decision-making (Prinsloo & Slade, 2014).

In face-to-face courses, instructors are more accustomed to observing and responding to educational data, relying on a plethora of both verbal and non-verbal cues, such as facial expressions or body language (Herodotou et al., 2019; Vanlommel et al., 2017). However, while teaching online, instructors have indirect interaction with students, and they are less exposed to non-verbal communication (Barak & Usher, 2020; Picciano, 2012). In this scenario, many of the learners' actions and behavior (such as navigating through the course pages) might be harder to track (Picciano, 2012; Siemens, 2013). This situation compromises the data-driven instructional process since the instructors' rely on only a thin layer of easily observed students' course activity (Alexandron et al., 2019; Gašević et al., 2016).

Even so, web-based environments may assist in the data-driven instructional process, as they usually keep digital traces left by learners as they engage with the course material and assignments (Vigentini et al., 2017). Over time, the value of such digital traces has been recognized as a promising source of data about students' learning processes (Gašević et al., 2016; Kim et al., 2016). The application of such data methods in higher education settings helped to

inform the rise of the field of learning analytics (Larrabee S nderlund et al., 2019; Siemens, 2013).

According to the Society for Learning Analytics for Research (SOLAR), learning analytics (LA) refer to the measurement, collection, analysis and reporting of data about learners and their contexts in order to understand and promote learning processes (Siemens, 2013). Higher education institutions are increasingly turning to LA tools to evaluate learners' online behavior and to analyze and interpret it to gain new insights (Larrabee S nderlund et al., 2019; Siemens, 2013).

LA are known for their great potential for improving teaching and learning in higher education contexts (Hilliger et al., 2020; Shibani et al., 2020). The bulk of empirical LA studies have focused on student performance measure tools, using LA systems for modeling and classifying learners' individual and collective needs, behavior, and performances (Baig et al., 2020; Muljana & Luo, 2020). So why is their institutional adoption still limited?

Several recent studies provide possible explanations for this question. First, it seems that LA designers and researchers tend to disconnect the main stakeholders from the design process, meaning, they create the analytics tools without focusing specifically on users' preferences, needs, capabilities, and data demands (Holstein et al., 2019; Shibani et al., 2020). Therefore, it is not surprising that according to instructors' reports, a common challenge they are facing while implementing LA tools in their classes is that the data provided to them is not aligned with their needs and pedagogical perspectives (Klein et al., 2019; Shibani et al., 2020).

Second, raw data alone do not hold much meaning for the instructors; they must hold the ability to transform this data into information that can be useful for making decisions (Ruip rez-Valiente et al., 2017). Yet, not all instructors possess the data literacy skills necessary to produce meaning from the data (Hilliger et al., 2020; Klein et al., 2019). Data literacy is defined as the ability to understand and use data in an effective way that will yield informed educational decisions (Green et al., 2016; Mandinach & Gummer, 2013). Lacking the skills and knowledge needed to become data literate, along with the disconnection from the design process of LA tools, makes it difficult for the instructors to derive meaningful insights from the data with the purpose of making informed decisions (Maisarah et al., 2021).

Most of the studies on instructors' use of LA tools to inform decisions have been undertaken in the context of face-to-face courses that use online learning management systems or in hybrid courses that combine face-to-face with online instruction. Yet, there is a lack of studies that specifically address the perspectives of instructors who teach in massive open online courses (MOOCs).

Data-Driven Instruction in Massive Open Online Courses

Over the last two decades, online education has seen the emergence and adoption of MOOCs. MOOCs are web-based learning environments, designed to provide free and accessible high-quality education to the masses (Dillahunt et al., 2014; Mcauley et al., 2010). Since they were first introduced in 2008, the popularity of MOOCs has been growing rapidly among learners and researchers worldwide (Meek et al., 2017; Zhu et al., 2020). At the end of 2020, more than sixteen thousand MOOCs were offered by more than 950 universities with more than 180 million enrollees (Shah, 2020). MOOC popularity has reached new levels due to the outbreak of the COVID-19 pandemic. The top three MOOC providers (i.e., Coursera, edX, and FutureLearn) registered as many new users in April 2020 as in the whole of 2019.

While sharing many characteristics with online courses, MOOCs are somewhat different (Barak & Usher, 2022). Among other things, MOOCs allow a massive participation, with some MOOCs attract up to tens or even hundreds of thousands of learners, who come from a wider range of cultures and backgrounds than in typical college courses (Barak & Usher, 2022; Kizilcec & Brooks, 2017; Meek et al., 2017). With millions of global learners from various backgrounds, and with modern online platforms that enable the collection of fine-grained data of learners' activity and behavior, MOOCs offer a unique opportunity for data-driven educational research (Drachsler & Kalz, 2016; Kizilcec & Brooks, 2017).

In accordance, in the recent years, a significant body of research has examined the use of LA tools that provide educational data in the context of MOOCs (Drachsler & Kalz, 2016; Rizvi et al., 2020; Romero & Ventura, 2017). As recent meta-analyses show, the bulk of studies about data use in online courses, and in MOOCs in particular, focused on learners' needs, behavior, and performances (Baig et al., 2020; Romero & Ventura, 2017; Zhu et al., 2020). Although numerous LA models and techniques have been developed for MOOC settings, their implementation in real-world contexts remains limited (Er et al., 2019).

One main reason for the limited impact is that only a few studies address the perspectives of MOOC instructors regarding the implementation of such techniques in their courses (Er et al., 2019; Ga ević et al., 2016). Understanding MOOC instructors' perspective regarding data use is critical since they are important stakeholders in the process of adopting and implementing innovative learning technologies (Siemens, 2013; Zhu et al., 2020). MOOC instructors are the ones who access and evaluate the educational data, draw conclusions, and take actions to support students and improve their courses (Alexandron et al., 2019; Leitner et al., 2017).

Recently, studies have begun examining instructors' perspectives about the use of educational data in face-to-face courses that use online learning management systems or in hybrid courses (Hilliger et al., 2020; Klein et al., 2019; Shibani et al., 2020). However, given that MOOCs are non-formal learning settings with massive numbers of learners from diverse backgrounds (Kizilcec & Brooks, 2017), previous findings from formal learning contexts should not be taken for granted (Er et al., 2019). Yet, there is a lack of studies that provide qualitative evidence about interest in educational data, and barriers to using them, from the perspective of higher education instructors who teach in MOOCs. The current study aims at bridging this gap.

Methods

Research Participants

Our participants included 12 higher education instructors (7 males, 5 females) from four public universities in Israel, each teaching in a different MOOC. All MOOCs were in science, technology, engineering, and mathematics (STEM) domains (6 from the applied sciences, 4 from the natural sciences, and 2 from medicine). The participants reported on having sufficient prior experience in using technology in the classroom. They were 40–70 years old ($M=51.3$, $SD=8.6$), with 5–40 years of teaching experience in higher education ($M=20.8$, $SD=11.1$), and 1–5 years of teaching experience in MOOCs ($M=3.2$, $SD=9.1$). To recruit participants, we contacted potential respondents, from the authors' professional and personal networks, via email, and continued recruitment using snowball sampling (Goodman, 1961). To protect the instructors' anonymity, names and identifying details were concealed by codes.

Research Methods and Tools

The study applied a qualitative phenomenological research design, in which the researchers describe the lived experiences of individuals about a phenomenon as described by the participants (Creswell, 2014). As customary in such studies, data collection involved in-depth semi-structured interviews, which were conducted during February–April 2020. The interviews took place via synchronous technologies, such as Skype¹ or Zoom,² and lasted approximately 60 min each. The interviews were audio recorded and fully transcribed before analysis. The interview sessions included seven key questions that focused on instructors' interest in

educational data and perceived barriers to using them for decision-making, as follows:

1. Please describe the MOOC you teach and your prior experience in online teaching in general and in MOOCs in particular.
2. Could you elaborate on the types of data about MOOC learners' activity that interest you the most?
3. Whether and how do you keep track of the data that interest you about learners' activity?
4. Could you elaborate on the main sources from which you collect such data?
5. In which ways do you prefer the data to be presented to you via the MOOC platform?
6. Did you encounter any specific challenges/barriers while trying to monitor or engage with data about learners' activity? If so, please elaborate.
7. Would you consider taking actions/making changes to improve your MOOC based on the data about learners' activity? If so, please elaborate.

Data Analysis

The interview data were qualitatively analyzed by the first author using the conventional (inductive) data analysis approach, which resulted in the establishment of a comprehensive set of themes. To ensure inter-coder reliability, a sample of 2 full interview transcripts, along with the established set of themes, was sent to the second author. The inter-rater agreement between the two authors was calculated using Cohen's kappa analysis, indicating good reliability of 0.83. Also, as a result of this comparative exercise, few themes were renamed to enhance comprehension and others were merged to avoid overlaps. The main themes identified from the data addressing each research question are discussed in the following sections.

Findings

MOOC Instructors' Interest in Educational Data

The analysis has raised two categories for instructors' interest in educational data.

Social Interactions

The first category refers to the instructors' interest in data about learners' social interactions with their peers while taking the MOOC. Out of the 12 interviewees, five mentioned that they monitored students' social learning experiences by following messages posted by students on the discussion boards. As one instructor put it, "I believe learners have

¹ <https://www.skype.com/en/>.

² <https://zoom.us/>.

a desire to be a part of a learning community and that we should create such opportunities for them. I make sure to monitor the interactions between learners as much as I can, mainly through the MOOC forums” (I3, male). He further expressed his feelings that the MOOC platform does not make this type of data accessible enough to the instructors: “I find myself spending long hours trying to track students’ correspondence in the forum [...] trying to figure out who those students are. It did not feel like these data were accessible enough to me as a MOOC instructor.”

Another instructor noted that what he found most interesting was that the students in his MOOC turned to the discussion board mostly to receive educational support from their peers:

Once a week I enter the forum designated for students to consult with each other. Students ask other students what they understood from the lecture [...] I find it fascinating to read the answers students provide to each other. The way they really try to help and support each other. (I4, male)

This instructor also mentioned that the data is not accessible enough for the instructors, which made him “[...] spending a lot of time and effort searching for it and cross-referencing it with other aspects, such as the demographic background of the learner who started the thread” (I4, male).

The instructors’ interest in learners’ social interactions was sometimes linked directly to the lack of face-to-face interaction between MOOC learners. As one interviewee put it, MOOC instructors should express interest in the social aspects of taking a MOOC since it is important to “know what it feels like to be a part of a pretty isolated learning community, to not know who your classmates are and not meeting them in-person” (I8, male).

Several instructors noted the ways in which they took actions aimed at bridging the lack of social interactions in MOOCs. One instructor shared that she opened a designated thread for learners to introduce themselves and get to know each other:

Learners were eager to introduce themselves to their classmates. Our goal was to encourage a social collaboration between students and to reduce the lack of interactivity and feeling of isolation students might be facing. (I11, female)

Yet, she shared that she found it difficult to keep track of all the data from students’ correspondence and expressed her desire that “such information would be available to us in an interesting and informative manner.”

Another instructor stated she tried to “minimize the distance between the learners” by encouraging them to perform the final assignment in groups: “We agreed to include a final assignment in which students should form a group of

4 and work together on a shared PowerPoint presentation” (I7, female).

Educational Resources

The second category refers to the instructors’ interest in data about problems and issues students are having while engaging with the MOOC educational resources. The most mentioned one was related to course assignments that were evaluated by fellow students (i.e., peer assessment assignments), which was mentioned by seven out of the 12 instructors.

One instructor stated that the peer assessment activity in her course posed some challenges for the learners “who claimed that the peer assessment they received was done in a non-professional and unreliable manner.” According to her, this is an essential data that she must be aware to as close as possible to real time so she could “try to find a solution for those students” (I11, female). Another instructor also mentioned his desire to be aware of reliability issues with the peer assessment activities, since it might “cause students to drop out of the MOOC” (I6, male).

Interestingly, each of these instructors made a different decision to solve this problem. The first instructor shared that “several times the course staff and I read students’ projects and performed the assessment ourselves” (I11, female), while the second instructor revealed that he “decided to remove the open assignment and now the course includes only quizzes” (I6, male).

Another issue with the course resources that was repeatedly mentioned by the instructors was related to course assignments that were conducted in small groups (i.e., team assignments). This issue was mentioned by four out of the 12 instructors. Most instructors expressed their desire to be aware of problems students are facing while working on team assignments. For example, one instructor referred to the difficulty faced by the learners to find teammates to collaborate with:

The thing that interests me the most, or that I feel I should be informed about, is probably problems that learners encounter during the MOOC. The best example is the many posts uploaded to the online discussion board written by students who are searching for peers to collaborate with for the group project. (I8, male)

When asked whether he would consider taking actions based on the exposure to data about such difficulty, the instructor answered that he did approve for some specific students who contacted him to perform the team project individually. He explained his decision as follows: “[...] if I had known that many learners are struggling with the group project, I would have considered making some changes in the guidelines so that there would be an alternative for those

who fail to work effectively in a group with people from remote locations” (I8, male).

MOOC Instructors’ Barriers to Data Use for Decision-making

The analysis has raised four main barriers for using educational data for decision-making.

Lack of Customized Data

The first barrier mentioned by the instructors was the lack of data that is tailored to their personal needs, rather than according to general pre-determined sections. Five out of the 12 interviewees addressed the problem of sending the same types of data to the entire population of MOOC instructors, stating that MOOC providers should be cautious of the “one-size-fits-all” approach, as can be seen in the below quote:

It is very likely that what interests me and what interest the rest of the lecturers in our team are completely different things and the people in charge should take that to consideration [...] I would like to know about the age range of those students who paid for the MOOC, and my colleague is interested in their academic background. In the current version and as the data is currently presented to us, we cannot see a clear breakdown by the parameters that are of interest to each of us, or it is not available enough. (I3, male)

Another instructor articulated the same sentiment stating that MOOC providers should consult the instructors regarding the types of data that interest them: “I never received an email from the MOOC platform providers in which they ask me personally which data might interest me, they just send me what they think is of interest to everyone” (I11, female).

Several instructors have clearly raised the concern that receiving data that is not tailored to their personal needs leads them not to take actions to improve the courses, as one of the interviewees put it: “I would have considered taking actions if only I had access to more meaningful information. When I approach the instructor-facing dashboard, the most prominent data is the number of enrollees and the countries from which they come. Why? Who made the decision that this is the information that most lecturers want or need? These are not necessarily the data that interests me” (I12, female).

Finally, the need to periodically change the customized data was mentioned by one of the interviewees: “[I would like] a system where each of us can personally define what kind of data interests him [...] I think it is important that the lecturer would be able to occasionally change the settings of the data he would like to be alerted about and according to

the data he would make the necessary decisions to upgrade the course” (I1, male).

Lack of Real-time Access

The second barrier mentioned by the instructors was the lack of access to data immediately as they become available, or very rapidly after. Seven out of the 12 interviewees addressed the problem of receiving data in delay, stating that “the data usually come to my attention only after a while” (I1, male). Another instructor even stated that the lack of access to data immediately as they become available even affects his level of interest: “knowing that I receive information about my students’ activity only sometime after it happens, somehow makes me less interested” (I3, male).

Notably, five interviewees revealed they often find themselves refraining from taking actions to help their students and/or improve their MOOCs due to the lack of real-time data. According to one instructor, not having real-time access to learners’ data may affect her ability to intervene if necessary. She stated that it is important that the data would be available to her in the actual moment when things are happening since this way: “in case students are unsatisfied with some aspect of the course, I would be able to intervene and try to help” (I7, female).

A second instructor noted that lack of tutor intervention due to lack of access to real-time data could lead to student dropout: “I think that there were students who dropped out of the course because of difficulties they were facing with the assignments and the fact that they did not receive support when they needed it” (I6, male).

The next quote reveals a situation in which the instructor chose not to take any action since she did not receive information about the problem students were facing immediately when it happened:

There have been instances where students wrote inappropriate messages to each other on the course discussion board. The problem was that several times I found out about this only a while after it happened. This created a situation where I felt that the unpleasant discourse had already calmed down, so I chose not to take any action about it. (I10, female)

Lack of Data Literacy

The third barrier mentioned by the instructors was the lack of ability to understand and use data effectively to inform decisions. Half of the interviewees (6 of 12) described how the large data sets available confuse them and “make it difficult to monitor learners’ data and to derive meaning from the data” (I12, female).

Several instructors specifically mentioned the difficulty to analyze large data sets from multiple sources that are not combined into a unified and coherent file. As articulated by one of the instructors, various types of data that are of interest for her may indeed be available, but integrating them to produce meaning is not intuitive:

I have data about the total percentage of students who watched the videos, but I do not know enough about those students – at what point did they quit? What is their academic background? I mean, I have this information somewhere, but every piece of information is somewhere else on the platform, and it is almost impossible to integrate between all those different layers of data. (I2, female)

Another example can be seen from the following statement:

I have access to some interesting data [...] for example the overall number of participants divided according to their gender, age, continents, and academic disciplines. Right now, each set of data is presented in a different figure, and each figure is presented separately, which makes it very difficult to keep track of.” (I6, male).

The need for integrated data is particularly important for those instructors who lack experience and knowledge in data use: “The main thing is not the data, but the ability to ask the right questions. This is a complex and difficult skill to develop that requires a tremendous effort and valuable time [...] Most of them [MOOC instructors] may have a great interest in the data, but they lack the understanding, and the ability to know what and where to look. If the access to the data was simpler, then they might have considered using it more” (I5, male).

Lack of Institutional Support

The fourth barrier mentioned by the instructors was the lack of sufficient support from the academic institution in which they teach. More than half of the interviewees (7 of 12) stated that the “lack of guidance and prior preparation” prevented them from tracking learners’ data in a more rigorous and routine manner.

Several instructors indicated that they did not have enough time to engage with the data even though they found it important for the improvement of their MOOCs, as one instructor stated: “it takes so much time to search, analyze, and make meaning from the data [...] Time is actually my biggest problem when I think about monitoring data about my MOOC learners” (I8, male). Another interviewee added that “we want to keep tracking the data about our learners and obviously, there is still much to improve. But our busy schedule requires each of us to invest in our own research” (I9, male). A central aspect that emerges from the quotes above is the instructors’ concern that engaging with the data would come

Table 1 Themes emerging from thematic analysis

Main theme	Subthemes
Interest in educational data	<ul style="list-style-type: none"> •Social interactions •Educational resources
Perceived barriers for data use	<ul style="list-style-type: none"> •Lack of customized data •Lack of real-time access •Lack of data literacy •Lack of institutional support

at the expense of their academic career, which will not be positively accepted by the university decision-makers. The instructors felt that after the initial investment of developing the MOOC, they were expected to go back to their academic obligations, rather than continue investing in the MOOCs:

What did I get from the MOOC as a faculty member? Or as an academic researcher? [...] Instead of developing the MOOC and keep working on it to improve it, I could have published at least two academic papers. (I1, male)

Lastly, three interviewees pointed out the need for dedicated workshops for using educational data efficiently to make informed decisions while teaching a MOOC. This could be clearly understood from the next quote:

The universities do not seem to understand the importance of providing us with technical and pedagogical support for using data [...] this is an expertise [...] people are not necessarily born with this skill; it has to be taught. (I8, male)

To summarize, the analysis of the interview data identified 6 key themes (see Table 1) that are aligned according to the two research questions: MOOC instructors’ interest in educational data and their perceived barriers to data use for decision-making.

Discussion

This study extends the scope of previous research by providing qualitative evidence about interest in educational data, and barriers to data use, from the perspective of higher education instructors who teach in MOOCs.

As our findings suggest, the participating instructors mostly showed interest in two main types of data. The first is data about social interactions that occurred between the remote learners. The instructors expressed their frustration that this type of data is not accessible to them in an easy and understandable way. Collecting data about social interactions was expressed by them as a challenging and time-consuming task. This finding validates previous studies, which reported that the implementation of data-driven instruction took

a lot of time and effort on behalf of the instructors, which increased their already stressed schedule (Klein et al., 2019; Shibani et al., 2020; Tsai et al., 2018).

The instructors' interest in learners' social interactions was sometimes linked to the lack of face-to-face interaction between MOOC learners. Several instructors expressed concern that MOOC learners might feel socially isolated and described their attempts to bridge this isolation by providing learners with opportunities to interact and collaborate. This finding provides empirical evidence that complements previous studies on the prominent role of social collaborations in MOOCs, where interpersonal interactions among learners are limited (Fang et al., 2019; Watted & Barak, 2018). These challenges may be even harder to deal with in times of emergency, such as the current COVID-19 pandemic, when extreme measures like quarantine or lockdown are taken, thereby potentially increasing a sense of loneliness (Murphy, 2020). Hence, the instructors in our study might have been concerned about their students' social state, and thus were inclined to take actions aimed at bridging the lack of social interactions.

The instructors further showed great interest in data about problems students are having while engaging with the MOOC educational resources, mainly the peer assessment assignments. Engaging with peer assessment activities in MOOCs was reported as associated with lower student satisfaction and with greater attrition (Jordan, 2015; Kulkarni et al., 2013), which may have affected the instructors' tendency to show interest in such data.

Notably, the participating instructors hardly mentioned a desire in data that relate to students' achievements and learning outcomes. Almost no mention was made of interest in data about students that failed a quiz or did not submit an assignment. This is surprising given that these students' activities are widely documented in the literature that deals with the utilization of LA tools in MOOCs (e.g., Lu et al., 2017). This may fall into the gap between research and practice identified by several researchers, who claimed that there is a disconnection between LA design and users' needs and expectations (Klein et al., 2019; Shibani et al., 2020), which emphasizes the need in an in-depth analysis of instructors' perspectives.

With regard to the instructors' perceived barriers to using educational data for decision-making, four barriers were mostly mentioned. The instructors referred to the difficulty that the data is not customized to their personal needs and is not brought to them in real time. This is in line with the claim that the learning analytics field should move from the "one-size-fits-all" approach to an approach that better emphasizes different perspectives (Knight et al., 2016). These findings also corroborate previous studies that reported the lack of personalized and timely data as the main

challenges instructors are facing while implementing LA in their courses (Hilliger et al., 2020; Klein et al., 2019; Shibani et al., 2020), which may compromise their ability to properly interpret students' behavior and to intervene in time to assist them (Knight et al., 2016; Picciano, 2012).

Moreover, it was often mentioned by the instructors that they refrained from taking actions due to their inability to understand and use the data effectively. The instructors described the difficulty to analyze and derive meaning from the large data sets that were available to them. This result aligns with, and sheds more light on, previous studies that point out data literacy as one of the key competencies that citizens in general, and higher education instructors specifically, must master in the society of the future (e.g., Raffaghelli & Stewart, 2020). There is a strong need to promote data literacy skills among educators, which would allow them to efficiently evaluate teaching and learning processes to perform the necessary instructional changes (Green et al., 2016).

Lastly, the instructors revealed that they refrained from taking actions due to the lack of sufficient support from their academic institutions. Mostly, the instructors expressed concern that engaging with educational data would come at the expense of their academic research. This finding confirms the assumptions that MOOC development and instruction are not supported by higher education institutions to a satisfying extent (Lorenz, 2016). Several instructors specifically stressed their desire to receive additional guidance and support for properly engaging with educational data. This finding confirms the conclusions drawn from several studies about the importance of providing instructors with supporting mechanisms for the development of data-driven instruction expertise (Sakala & Chigona, 2020; Tsai et al., 2018).

Study Limitations

The first limitation derives from the study's qualitative nature, in which we relied on self-report measures (i.e., semi-structured interviews). The literature points out to several inherent deficiencies that might limit the generalizability of findings observed from self-reports. For instance, self-reports provide indirect information filtered through the views of interviewees and they might be biased by researcher's presence (Creswell, 2014). Still, previous research has posited that self-reports can be reliable under the fulfillment of several conditions: the information is known to respondents; the questions are phrased clearly and unambiguously, relate to recent activities, and require a serious and thoughtful response; and answering them will not lead to embarrassing or threatening disclosures (Dang et al., 2020; Kuh, 2002). We believe these conditions were met in the current study, and hence, the instructors' responses could be

considered valid findings. The second limitation relates to the research population and setting. Our findings were obtained from the perspective of 12 instructors who teach in four flagship universities in one country. Thus, suggestions for future work include expanding the research settings to other higher education institutions from different countries, to capture insights of engaging with data-driven information in MOOCs worldwide.

Summary and Future Research

MOOCs provide students with opportunities to communicate and interact with thousands of learners from different countries and to pose and solve problems collaboratively and cross-culturally (Barak & Usher, 2022). Indeed, it has been argued that due to their global reach and diverse nature, it is vital to understand whether MOOCs provide learners with sufficient opportunities to gain 21st century thinking skills (Cui et al., 2014; Zimmerman et al., 2017). According to Cui et al. (2014), MOOCs provide minimal supervision and support, hence, MOOC learners should be highly self-directed to attain their learning outcomes. They often master new knowledge by the use of self-assessment and peer assessment activities. Such learning processes were documented as promoting learners' collaborative learning, self-directed learning, and creative and innovative abilities (Barak & Usher, 2020, 2022; Zhu & Bonk, 2019). All of which are considered important skills for dealing with future challenges. From the instructors' perspective, limited research has been done on the many 21st century skills required to effectively design and teach in a MOOC (Zhu et al., 2020). Due to the technical structure of MOOCs, instructors can record educational data and track learners' learning process using a variety of technological tools. With the help of such data, instructors can better guide learners, promote collaborations among them and create a diverse community of learners (Cui et al., 2014). Yet, to effectively implement and use educational data to improve their courses, MOOC instructors should master a wide range of 21st century skills, such as analytical thinking - the ability to use and interpret large datasets to inform decisions (Hilliger et al., 2020; OECD, 2018; Raffaghelli & Stewart, 2020). Our findings showed that the types of educational data that MOOC instructors showed great interest in are not accessible to them in an easy, understandable and actionable way. In this regard, we believe that future research should focus on how to actively co-design with MOOC instructors technological tools that will support the data-driven instructional process, an essential step to ensure the alignment of such tools with real-world needs and their usefulness and usability in real-world educational contexts (An et al., 2020;

Holstein et al., 2019). It is also important for future research to examine the contribution of MOOCs to the acquisition of the required thinking skills that meet the challenges of the 21st century.

Funding This work has been supported by the “Applications of big data and artificial intelligence to online learning technologies” project, under research grant 0607018111 from the Israeli Ministry of Science and Technology.

Declarations

Ethical Statement This research was granted with an approval from the university's Ethics Committee.

Consent Statement To ensure the research is conducted in an ethical manner, the participants were informed about the research goal, process, and rights. The participants were informed that participation is voluntary, and they were given the choice to withdraw at any time. To protect their anonymity, names and contact information were concealed by codes.

Conflict of Interest The authors declare no competing interests.

References

- Alexandron, G., Yoo, L. Y., Ruipérez-Valiente, J. A., Lee, S., & Pritchard, D. E. (2019). Are MOOC learning analytics results trustworthy? With fake learners, they might not be! *International Journal of Artificial Intelligence in Education*, 29(4), 484–506. <https://doi.org/10.1007/s40593-019-00183-1>
- An, P., Holstein, K., D'Anjou, B., Eggen, B., & Bakker, S. (2020). The TA framework: Designing real-time teaching augmentation for K-12 classrooms. *Conference on Human Factors in Computing Systems - Proceedings*. <https://doi.org/10.1145/3313831.3376277>
- Baig, M. I., Shuib, L., & Yadegaridehkordi, E. (2020). Big data in education: A state of the art, limitations, and future research directions. *International Journal of Educational Technology in Higher Education*, 17(1). <https://doi.org/10.1186/s41239-020-00223-0>
- Barak, M. & Usher, M. (2020). Innovation in a MOOC: Project-based learning in the international context. In J. J. Mintzes & E. M. Walter (Eds.) *Active Learning in College Science: The Case for Evidence Based Practice*. Berlin: Springer Nature, pp. 639–653. https://doi.org/10.1007/978-3-030-33600-4_39
- Barak, M., & Usher, M. (2022). The innovation level of engineering students' team projects in hybrid and MOOC environments. *European Journal of Engineering Education*, 47(2), 299–313. <https://doi.org/10.1080/03043797.2021.1920889>
- Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches* (4th ed). SAGE Publications, Inc.
- Cui, L., Li, H., & Song, Q. (2014). Developing the ability for a deep approach to learning by students with the assistance of MOOCs. *World Transactions on Engineering and Technology Education*, 12(4), 685–689.
- Dang, J., King, K. M., & Inzlicht, M. (2020). Why are self-report and behavioral measures weakly correlated? *Trends in Cognitive Sciences*, 24(4), 267–269.
- Dillahunt, T., Wang, Z., & Teasley, S. D. (2014). Democratizing higher education: Exploring MOOC use among those who cannot afford a formal education. *The International Review of Research in Open*

- and Distance Learning, 15(5), 177–196. <https://doi.org/10.19173/irrodl.v15i5.1841>
- Drachler, H., & Kalz, M. (2016). The MOOC and learning analytics innovation cycle (MOLAC): A reflective summary of ongoing research and its challenges. *Journal of Computer Assisted Learning, 32*(3), 281–290. <https://doi.org/10.1111/jcal.12135>
- Er, E., Gómez-Sánchez, E., Dimitriadis, Y., Bote-Lorenzo, M. L., Asensio-Pérez, J. I., & Álvarez-Álvarez, S. (2019). Aligning learning design and learning analytics through instructor involvement: A MOOC case study. *Interactive Learning Environments, 27*(5–6), 685–698. <https://doi.org/10.1080/10494820.2019.1610455>
- Fang, J., Tang, L., Yang, J., & Peng, M. (2019). Social interaction in MOOCs: The mediating effects of immersive experience and psychological needs satisfaction. *Telematics and Informatics, 39*(August 2018), 75–91. <https://doi.org/10.1016/j.tele.2019.01.006>
- Gašević, D., Dawson, S., Pardo, A., Gašević, D., Dawson, S., & Pardo, A. (2016). How do we start? State and directions of learning analytics adoption. *2016 ICDE Presidents' Summit, December*, 1–24. <https://doi.org/10.13140/RG.2.2.10743.42401>
- Goodman, L. A. (1961). Snowball sampling. *Annals of Mathematical Statistics, 32*(1), 148–170. <https://doi.org/10.1214/AOMS/1177705148>
- Green, J. L., Schmitt-Wilson, S., Versland, T., Kelting-Gibson, L., & Nollmeyer, G. E. (2016). Teachers and data literacy: A blueprint for professional development to foster data driven decision making. *Journal of Continuing Education and Professional Development, January*. <https://doi.org/10.7726/jcepd.2016.1002>
- Herodotou, C., Hlosta, M., Borooa, A., Rienties, B., Zdrahal, Z., & Mangafa, C. (2019). Empowering online teachers through predictive learning analytics. *British Journal of Educational Technology, 50*(6), 3064–3079. <https://doi.org/10.1111/bjet.12853>
- Hilliger, I., Ortiz-Rojas, M., Pesántez-Cabrera, P., Scheihing, E., Tsai, Y. S., Muñoz-Merino, P. J., Broos, T., Whitelock-Wainwright, A., Gašević, D., & Pérez-Sanagustín, M. (2020). Towards learning analytics adoption: A mixed methods study of data-related practices and policies in Latin American universities. *British Journal of Educational Technology, 51*(4), 915–937. <https://doi.org/10.1111/bjet.12933>
- Holstein, K., McLaren, B. M., & Alevan, V. (2019). Co-designing a real-time classroom orchestration tool to support teacher–AI complementarity. *Journal of Learning Analytics, 6*(2), 27–52. <https://doi.org/10.18608/jla.2019.62.3>
- Jordan, K. (2015). Massive open online course completion rates revisited: Assessment, length and attrition. *International Review of Research in Open and Distance Learning, 16*(3), 3451–358. <https://doi.org/10.19173/irrodl.v16i3.2112>
- Kim, D., Park, Y., Yoon, M., & Jo, I. H. (2016). Toward evidence-based learning analytics: Using proxy variables to improve asynchronous online discussion environments. *Internet and Higher Education, 30*, 30–43. <https://doi.org/10.1016/j.iheduc.2016.03.002>
- Kizilcec, R. F., & Brooks, C. (2017). Diverse big data and randomized field experiments in MOOCs. *Handbook of Learning Analytics, 211–222*. <https://doi.org/10.18608/hla17.018>
- Klein, C., Lester, J., Rangwala, H., & Johri, A. (2019). Technological barriers and incentives to learning analytics adoption in higher education: Insights from users. *Journal of Computing in Higher Education, 31*(3), 604–625. <https://doi.org/10.1007/s12528-019-09210-5>
- Knight, D. B., Brozina, C., & Novoselich, B. (2016). An investigation of first-year engineering student and instructor perspectives of learning analytics approaches. *Journal of Learning Analytics, 3*(3), 215–238. <https://doi.org/10.18608/jla.2016.33.11>
- Kuh, G. D. (2002). The National Survey of Student Engagement: Conceptual framework and overview of psychometric properties. *Framework & Psychometric Properties, 1*(1), 1–26. <https://doi.org/10.5861/ijrse.2012.v1i1.19>
- Kulkarni, C., Wei, K. P., Le, H., Chia, D., Papadopoulos, K., Cheng, J., Koller, D., & Klemmer, S. R. (2013). Peer and self assessment in massive online classes. *ACM Transactions on Computer-Human Interaction, 20*(6). <https://doi.org/10.1145/2505057>
- Larrabee Sønderlund, A., Hughes, E., & Smith, J. (2019). The efficacy of learning analytics interventions in higher education: A systematic review. *British Journal of Educational Technology, 50*(5), 2594–2618. <https://doi.org/10.1111/bjet.12720>
- Leitner, P., Khalil, M., & Ebner, M. (2017). Learning analytics: Fundamentals, applications, and trends. *Learning Analytics: Fundamentals, Applications, and Trends, Studies in Systems, Decision and Control, 94*(February), 1–23. <https://doi.org/10.1007/978-3-319-52977-6>
- Lorenz, A. (2016). The MOOC production fellowship: Reviewing the first German MOOC funding program. In M. Khalil, M. Ebner, M. Kopp, A. Lorenz, & M. Kalz (Eds.), *The European Stakeholder Summit on Experiences and Best Practices in and around MOOCs* (pp. 185–196).
- Lu, O. H. T., Huang, J. C. H., Huang, A. Y. Q., & Yang, S. J. H. (2017). Applying learning analytics for improving students engagement and learning outcomes in an MOOCs enabled collaborative programming course. *Interactive Learning Environments, 25*(2), 220–234. <https://doi.org/10.1080/10494820.2016.1278391>
- Maisarah, N., Khuzairi, S., & Cob, Z. C. (2021). A preliminary model of learning analytics to explore data visualization on educator's satisfaction and academic performance in higher education. *Springer International Publishing*. <https://doi.org/10.1007/978-3-030-90235-3>
- Mandinach, E. B., & Gummer, E. S. (2013). A systemic view of implementing data literacy in educator preparation. *Educational Researcher, 42*(1), 30–37. <https://doi.org/10.3102/0013189X12459803>
- Margaryan, A., Bianco, M., & Littlejohn, A. (2015). Instructional quality of massive open online courses (MOOCs). *Computers & Education, 80*, 77–83.
- Mcauley, A. A., Stewart, B., Siemens, G., & Cormier, D. (2010). *The MOOC model for digital practice*.
- Meek, S. E. M., Blakemore, L., & Marks, L. (2017). Is peer review an appropriate form of assessment in a MOOC? Student participation and performance in formative peer review. *Assessment and Evaluation in Higher Education, 42*(6), 1000–1013. <https://doi.org/10.1080/02602938.2016.1221052>
- Muljana, P. S., & Luo, T. (2020). Utilizing learning analytics in course design: Voices from instructional designers in higher education. *Journal of Computing in Higher Education*. <https://doi.org/10.1007/s12528-020-09262-y>
- Murphy, M. P. A. (2020). COVID-19 and emergency eLearning: Consequences of the securitization of higher education for post-pandemic pedagogy. *Contemporary Security Policy, 41*(3), 492–505. <https://doi.org/10.1080/13523260.2020.1761749>
- OECD. (2018). *The future of education and skills: Education 2030*. <https://www.oecd.org/education/2030/E2030> Position Paper (05.04.2018).pdf
- Picciano, A. G. (2012). The evolution of big data and learning analytics in American higher education. *Journal of Asynchronous Learning Network, 16*(3), 9–20. <https://doi.org/10.24059/olj.v16i3.267>
- Prinsloo, P., & Slade, S. (2014). Educational triage in open distance learning: Walking a moral tightrope. *International Review of Research in Open and Distance Learning, 15*(4), 306–331. <https://doi.org/10.19173/irrodl.v15i4.1881>
- Raffaghelli, J. E., & Stewart, B. (2020). Centering complexity in 'educators' data literacy' to support future practices in faculty development: A systematic review of the literature. *Teaching in Higher*

- Education*, 25(4), 435–455. <https://doi.org/10.1080/13562517.2019.1696301>
- Rizvi, S., Rienties, B., Rogaten, J., & Kizilcec, R. F. (2020). Investigating variation in learning processes in a FutureLearn MOOC. *Journal of Computing in Higher Education*, 32(1), 162–181. <https://doi.org/10.1007/s12528-019-09231-0>
- Romero, C., & Ventura, S. (2017). Educational data science in massive open online courses. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 7(1). <https://doi.org/10.1002/widm.1187>
- Ruipérez-Valiente, J. A., Muñoz-Merino, P. J., Pijera Díaz, H. J., Ruiz, J. S., & Kloos, C. D. (2017). Evaluation of a learning analytics application for open edX platform. *Computer Science and Information Systems*, 14(1), 51–73. <https://doi.org/10.2298/CSIS160331043R>
- Sakala, L. C., & Chigona, W. (2020). How lecturers neutralize resistance to the implementation of learning management systems in higher education. *Journal of Computing in Higher Education*, 32(2), 365–388. <https://doi.org/10.1007/s12528-019-09238-7>
- Shah, D. (2020). *By the numbers: MOOCs in 2020*. Class-Central. <https://www.classcentral.com/report/mooc-stats-2020/>
- Shibani, A., Knight, S., & Buckingham Shum, S. (2020). Educator perspectives on learning analytics in classroom practice. *Internet and Higher Education*, 46(February), 100730. <https://doi.org/10.1016/j.iheduc.2020.100730>
- Siemens, G. (2013). Learning analytics: The emergence of a discipline. *American Behavioral Scientist*, 57(10), 1380–1400. <https://doi.org/10.1177/0002764213498851>
- Tsai, Y., & Hsun, Lin, C. hung, Hong, J. chao, & Tai, K. hsin. (2018). The effects of metacognition on online learning interest and continuance to learn with MOOCs. *Computers and Education*, 121, 18–29. <https://doi.org/10.1016/j.compedu.2018.02.011>
- Usher, M., Barak, M., & Haick, H. (2021a). Online vs. on-campus higher education: Exploring innovation in students' self-reports and students' learning products. *Thinking Skills and Creativity*, 42, 100965. <https://doi.org/10.1016/j.tsc.2021.100965>
- Usher, M., Hershkovitz, A., & Forkosh-Baruch, A. (2021b). From data to actions: Instructors' decision making based on learners' data in online emergency remote teaching. *British Journal of Educational Technology*, 52(4), 1338–1356. <https://doi.org/10.1111/bjet.13108>
- Vanlommel, K., Van Gasse, R., Vanhoof, J., & Van Petegem, P. (2017). Teachers' decision-making: Data based or intuition driven? *International Journal of Educational Research*, 83(March 1994), 75–83. <https://doi.org/10.1016/j.ijer.2017.02.013>
- Vigentini, L., Clayphan, A., & Chitsaz, M. (2017). Dynamic dashboard for educators and students in FutureLearn MOOCs: Experiences and insights. *CEUR Workshop Proceedings*, 1967(March 2017), 20–35.
- Wang, Q., & Woo, H. L. (2007). Comparing asynchronous online discussions and face-to-face discussions in a classroom setting. *British Journal of Educational Technology*, 38(2), 272–286. <https://doi.org/10.1111/j.1467-8535.2006.00621.x>
- Watted, A., & Barak, M. (2018). Motivating factors of MOOC completers: Comparing between university-affiliated students and general participants. *Internet and Higher Education*, 37(June 2017), 11–20. <https://doi.org/10.1016/j.iheduc.2017.12.001>
- Yulina, I. K., Permanasari, A., & Hernani, h., & Setiawan, W. (2019). Analytical thinking skill profile and perception of pre service chemistry teachers in analytical chemistry learning. *Journal of Physics: Conference Series*. <https://doi.org/10.1088/1742-6596/1157/4/042046>
- Zhu, M., Sari, A. R., & Lee, M. M. (2020). A comprehensive systematic review of MOOC research: Research techniques, topics, and trends from 2009 to 2019. *Educational Technology Research and Development*, 68(4), 1685–1710. <https://doi.org/10.1007/s11423-020-09798-x>
- Zimmerman, C., Dreisiebner, D., & Hoffer, E. (2017). Designing a MOOC to foster critical thinking and its application in business education. *International Journal for Business Education*, 157(1). <https://doi.org/10.30707/IJBE157.1.1648132890.935577>
- Zhu, M., & Bonk, C. J. (2019). Designing MOOCs to facilitate participant self-monitoring for self-directed learning. *Online Learning*, 23(4), 106–134. <https://doi.org/10.24059/olj.v23i4.2037>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.