

MIT Open Access Articles

Using multi-platform learning analytics to compare regional and global MOOC learning in the Arab world

The MIT Faculty has made this article openly available. ***Please share*** how this access benefits you. Your story matters.

As Published: 10.1016/J.COMPEDU.2019.103776

Publisher: Elsevier BV

Persistent URL: <https://hdl.handle.net/1721.1/135904>

Version: Original manuscript: author's manuscript prior to formal peer review

Terms of use: Creative Commons Attribution-NonCommercial-NoDerivs License



Using Multi-Platform Learning Analytics to Compare Regional and Global MOOC Learning in the Arab World

José A. Ruipérez-Valiente^a, Sherif Halawa^b, Rachel Slama^a, Justin Reich^a

^a*Massachusetts Institute of Technology, 77 Massachusetts Ave, Cambridge, 02139, USA*

^b*Edraak, 300 King Abdullah St, Amman, Jordan*

Abstract

Recent studies of massive open online courses (MOOCs) have focused on global providers such as edX, Coursera, and FutureLearn, with less attention to local initiatives that target regional learners. In this study we combine data from the main edX platform and one regional MOOC provider, Edraak in Jordan, to explore differences in learners' behavior and preferences. We find that regional provider Edraak attracts younger learners, more females and those with lower levels of education compared to global providers. Edraak learners value local courses because they cater to their interests and learning needs. We document our multi-platform learning analytics procedure, where we establish a common data format and script that enables an “apples-to-apples” comparison without exchanging data—a common privacy and data security concern. These findings suggest the potential of this methodological approach to study and learn from regional MOOC providers, particularly around the questions of equity and access in the global MOOC ecosystem.

Keywords: Learning Analytics, Massive Open Online Courses, Large Scale Analytics, Cultural Factors, Distance Learning, Equity

Email addresses: jruipere@mit.edu (José A. Ruipérez-Valiente), shalawa@qrf.org (Sherif Halawa), rslama@mit.edu (Rachel Slama), jreich@mit.edu (Justin Reich)

1. Introduction

New virtual learning environments generate extensive data about learner behavior, and the field of learning analytics has emerged with the goal of using this data to optimize individual learning and learning contexts [1]. In the past decade, massive open online courses (MOOCs) have offered higher education institutions (HEIs) a new platform for engaging learners around the world and for studying patterns in learning [2]. Especially in their early years, the largest providers, English-language MOOC providers described their mission in terms of “democratizing higher education” by creating free, global, online access to courses from elite universities [3]. However, emerging evidence suggests that MOOCs, and global providers in particular, have not met this goal but have primarily reached educated learners from affluent countries [4], potentially widening educational disparities [5]. For example among global providers like Coursera, edX or FutureLearn, there are gaps in participation and persistence among females in STEM, learners with low levels of educational attainment and from less-developed countries [6, 4, 7]. For learners with low levels of English proficiency, linguistic and cultural factors further compound the challenges of online learning [8, 9, 10]. Thus, the need for scalable solutions to reach linguistically and culturally diverse learners with the most limited access to higher education opportunities persists.

English language MOOC providers that have aimed to attract global audiences are complemented by a set of regional MOOC providers that have been emerging over the past years with a focus on serving learners in their regions by developing localized content. Some examples include the Ibero-American initiative MiriadaX ², the Chinese initiative XuetangX ³, the Italian Fed-

²<https://miriadax.net/en/>

³<https://www.xuetangx.com/>

erica.EU ⁴, the German OpenHPI ⁵, the French platform France Université Numérique (FUN ⁶) or the Arab initiative Edraak ⁷. Several of these initiatives (Edraak, XuentangX and FUN) use Open edX ⁸—edX’s collaborative open source initiative—as the underlying software to run MOOCs. By focusing on
30 more specific socio-linguistic communities with a shared language and culture, these initiatives may be able to play a complementary role to MOOCs aiming at a global, primarily English-language audience.

Numerous studies have pointed to the importance of considering cultural factors when designing inclusive online learning experiences [8, 9, 10]. Yet, at
35 present, no studies have simultaneously capitalized on large-scale data from local and global MOOC providers – where course features such as language of instruction and forums, cultural background of instructor, and localized course examples can provide insight into learner behavior across diverse contexts. We present a case study of a multi-platform analysis on MOOCs using data from
40 three different institutions—HarvardX, MITx, and the regional Jordan-based MOOC platform, Edraak. MITx and HarvardX offer MOOCs via edX and Edraak hosts its own Open edX platform. MITx and HarvardX offer courses mainly in English that target a global audience while Edraak has localized its instance to focus on the Arab States, offering courses mainly in Arabic. We ex-
45 amine MOOC participation and engagement in the Arab world, by comparing differences between the global setting of MITx and HarvardX and the regional provider Edraak. Our methodology entails establishing a common data format and having each researcher execute a common analysis script in their local environment, bypassing two main issues related to cross-platform research – having
50 a unified and comparable dataset and data exchange. The proposed model of innovation diffusion for learning analytics in higher education can help us

⁴<https://federica.eu/>

⁵<https://open.hpi.de/>

⁶<https://www.fun-mooc.fr/>

⁷<https://www.edraak.org/>

⁸<https://open.edx.org/>

understand influential factors in online learning for diverse populations.

Our study capitalizes on the multi-platform, observational data from regional and global MOOC providers to better understand differences in learners' behavior across the two contexts. We explore the following three research questions:

- Do regional MOOC providers help narrow educational and gender gaps for the local population?
- Do regional MOOC providers have higher levels of engagement, participation and completion for the local population, compared to global providers?
- To what degree are course level factors (course topic, localized production of content, language of instruction) and self-reported individual learners' characteristics related to these outcomes?

The remainder of the paper is organized with a review of select research studies to contextualize our research questions in Section 2, a description of the methodology in Section 3, a presentation of the results and discussion in Sections 4 and 5, and we offer conclusions in Section 6.

2. Background

2.1. Learning Analytics

As more and more educational data becomes available, learning analytics researchers are playing a pivotal role in understanding variation in learning across contexts. One of the major challenges in the field has been to build the necessary conditions to implement learning analytics at an institutional level. Project SHEILA (Supporting Higher Education to Integrate Learning Analytics) represents one such effort [11]. SHEILA researchers worked between 2016 and 2018 towards building a policy development framework to support systematic, sustainable and responsible adoption of learning analytics at the institutional level.

Buckingham Shum provides some context for thinking about learning ana-
lytics across the entire system, introducing three levels where learning ana-
lytics can have an impact [12]. Macro-level analytics enables cross-institutional anal-
ysis, for example through institutional surveys or statewide access to individual
student data across a learner’s entire academic trajectory. Meso-level operates
at the institutional level such as in academic analytics, while micro-level analyses
focuses on supporting and tracking process-level data for individual learners and
groups of learner such as tracking online activity in clicks, location, and other
interpersonal data. Drachsler and Kalz mapped those levels to the MOOC and
Learning Analytics Cycle (MOLAC) where the micro level concerns teachers and
students in one course, the meso level informs a collection of MOOCs and their
course managers, and the macro attends to directory of MOOCs and informs
a whole community [13]. We situate our study– an examination of global and
regional learning analytics patterns across two platforms– at the macro level,
consistent with Drachsler and Kalz’s definition of cross-institutional learning
analytics as occurring at the macro level [13].

Multi-platform analytics are set apart from previous learning analytics stud-
ies. Prior studies have focused on a detailed analysis of one or few MOOCs (e.g.
[14]), a longitudinal analysis with many courses from one single MOOC provider
(e.g. [4, 15]), or a literature review of learning analytics in MOOCs (e.g. [16]).
Small MOOC studies have limited generalizability, while longitudinal studies
from one single institution do not capture variation across MOOC providers.
Literature reviews in this field have been unable to effectively make cross-study
comparisons due to the different methodologies employed. In response to these
challenges, we propose multi-platform learning analytics, which replicates the
same analyses using a common data format and script in different learning envi-
ronments, and we apply it to the MOOC context. This work aligns with calls for
methodologies that promote open education science [17] and replication stud-
ies in the social sciences [18]. Another study that has similar objectives is the
MOOC Replication Framework or MORF [19], a platform that enables institu-
tions to securely deposit their MOOC data and allows researchers to execute

110 Docker containers for data analysis while maintaining full privacy of the data.

One motivation for this study is the now well-documented finding over the past five years that global MOOC providers were most successful at reaching younger (ages 25-40), highly educated learners from more developed countries [20] and those learners have persisted and completed MOOCs at higher rates
115 than their peers from less developed countries [21, 5, 22]. Some of these trends may be explained by the challenges that are inherent in online learning at scale. Challenges faced by all students in online learning are even more salient for vulnerable participants in MOOCs. In addition to effective learning skills, successful students must have a high capacity for self-directed learning [23]. For
120 learners in some less developed countries, there are real concerns related to connectivity and access to a virtual platform. Pedagogical approaches that may be quite different than traditional learning environments in the participant's home country may also make the vulnerable learner's experience in the global MOOC challenging [24]. Additionally, some studies have raised particular concerns
125 about the Arab population regarding online learning, for example pointing out the numerous challenges in e-learning [25] and finding participation and completion gaps in MOOCs [22]. We believe that through looking at differences in learning analytics for a vulnerable population like Arab learners across two different MOOC platforms and contexts, we can shed some light on differences
130 in regional and global MOOC learning.

2.2. Language and Culture in Learning

In residential k-12 U.S. settings, longitudinal studies have demonstrated that learning in a second language may affect a learner's ability to benefit from content instruction in English and their ability to demonstrate knowledge and
135 skills on academic assessments in English [26]. However, many second language learners, particularly those with high levels of literacy in their home language [27], may adapt quickly and thrive academically in an English instructional setting.

In MOOCs, the understanding of the role of language and relatedly, culture

140 in learners' outcomes is emerging. One study [28] defines culture as "a complex and multi-dimensional construct that represents the shared values, beliefs, and basic assumptions of groups of people. It includes elements such as language, customs, social behaviour, and religion, and it influences how individuals relate to the world (p.27)". MOOC researchers have studied the link between individual learners' country of origin, language, and culture and various outcomes including MOOC persistence, completion, course and forum activity. Culture plays a role in off-task behaviors [29, 30], help-seeking [29], and collaboration [31, 29]. Liu *et al.*, [32] examined the impact of cultural and geographic factors on student performance in MOOC participants' course activity profiles, quiz activity profiles, and most connected forum peer or "best friends". They found that MOOC learners are more likely to interact with forum peers or "best friends" within their own cultural group which the authors suggest may signal students' common goals, a desire to eliminate language barriers through communication with someone of the same language group, or possibly a reflection of the students' offline relationships (e.g. they had taken the course together) [32].

160 Researchers have used various theoretical frameworks to interpret trends in learners' MOOC behavior on the basis of culture, including Hofstede's Cultural Dimensions Theory [33] which defines seven dimensions of culture based on international workforce survey data from the company IBM, the Cultural Dimensions of Learning Framework [34], which establishes eight cultural parameters related to social relationships, beliefs, and perceptions, [35] on understanding cultural differences, and hybrids of multiple models [32]. These frameworks provide useful lenses through which to interpret trends in MOOC learner behavior from the researcher perspective. The present case study documents a multi-platform learning analytics collaboration that leverages regional providers' in-depth of knowledge of local trends, previous work and the self-reported learner's survey results to incorporate the learner's perspective into the contextualization of the findings.

170 **3. Method**

This section describes the methodology in the context of MOOC providers Edraak and edX in Subsection 3.1, a description of the dataset, sample and measures in Subsections 3.2, 3.4 and 3.3 respectively. We end with an overview of the multi-platform MOOC learning analysis methodology in Subsection 3.5.

175 *3.1. Context*

In this work, we leverage the platform commonalities between Edraak and edX (with a focus on MITx and HarvardX courses). First, both platforms were founded with an aim to spread open education around the world, with edX targeting a global, English-speaking population and Edraak focusing on learners
180 in Arab states. Second, both institutions use Open edX software as their Content and Learning Management Systems (CMS and LMS) to create and teach MOOCs. The latter is particularly important in the context of multiplatform MOOC analytics as it facilitates the creation of a common data pipeline in the same format. Learners across both platforms will be interacting with courses in
185 an Open edX-based LMS, which facilitates the identification of trends in learner behavior across the platforms.

EdX was founded in 2012 by the Massachusetts Institute of Technology (MIT) and Harvard University. More than 100 schools, companies and other institutions have joined this initiative as partners to teach free courses to a
190 global audience. Open edX, which is the software used to run the edX main site, was created as an open source collaborative initiative to involve other educational institutions and developers around the globe. Open edX is being used by numerous institutions around the world to deliver their open or private courses, and according to their website, they have reached 40 million learners,
195 deployed over 20 thousand courses and support 32 languages. Hence, both edX and Open edX brands have reached significant numbers of learners worldwide.

Edraak was founded in 2013 by the Queen Rania Foundation for Education and Development. Edraak was designed to respond to the reported low levels

of English proficiency among Arab learners. One of the original missions was
200 to provide high-quality educational content in Arabic to fulfill these learners’
necessities (though some courses are also available in English). Edraak pro-
duces all of its courses in Arabic, except for courses teaching foreign languages,
and hosts them on its locally-adapted Open edX platform. Edraak’s courses
span multiple categories, including STEM, business and workforce development
205 skills, health, arts, and language. Course content is designed in collaboration
with regional experts from academia and industry. Edraak performs all other
course development and management tasks internally, such as media production,
content upload, course marketing, and course operations.

3.2. Dataset

210 We collected all MOOC data from MITx and HarvardX from 2012 to May
2018, and all data from Edraak from June 2014 to September 2018. The data
from each provider include self-reported variables by students from site registra-
tion and pre-course surveys and the Open edX tracking logs, that contain every
action performed in each course by each account. Part of the multi-platform
215 collaboration methodology that we establish for this case study involves having
a common data format for the log data, as we explain in Subsection 3.5.

MITx and HarvardX data contain 565 MOOC iterations, 12.67 million course
registrations from over 5.63 million unique student accounts, and we are able
to detect country of origin of 4.48 million accounts. Most of the courses taught
220 in MITx and HarvardX are in English. Edraak’s data contain 231 MOOC
iterations with a total of 3.77 million registrations to these courses from 1.5
million unique student accounts. We are able to identify the country of origin
of 1.48 million of these student accounts.

We also include self-reported survey results from both edX and Edraak. We
225 use pre-course surveys from MITx/HarvardX that ask learners to report their
experience with social identity threat and their level of English proficiency. We
use an engagement survey from Edraak participants to understand learners’
perceptions and characteristics. In Appendix A, we provide a description of the

surveys and relevant survey items.

230 3.3. Measures

Our multi-platform analytic approach involves defining a common person-course dataset where each row represents one registration in a course for one learner. The input fields have the following columns:

- User id: A unique identifier for each user account.
- 235 • Course id: A unique identifier for each course iteration in the course (one course can be run multiple times).
- Nationality: Country of origin of each account computed based on the modal IP address from the tracking logs of the student.
- Level of education: Qualitative variable with the self-reported level of
240 education of the account.
- Gender: Qualitative variable with the self-reported gender identification of the account.
- Viewed: Boolean variable that indicates if the account accessed the course at least once.
- 245 • Explored: Boolean variable that indicates if the account viewed at least half of the chapters of the course.
- Completed: Boolean variable that indicates if the account achieved a passing grade in the course.

We also use the Human Development Index (HDI) provided by the
250 United Nations (UN) in 2018 to summarize the development of a country [36].

Additionally, we categorize courses into four disciplinary categories defined by previous work [4]: Government, Health, and Social Science (GHSS), Humanities, History, Design, Religion, and Education (HHRDE), Computer Science (CS) and Science, Technology, Engineering, and Mathematics (STEM). The

255 distribution of the number of courses in each category in each platform is as follows:

- Edraak (N=231): 31.17% GHSS, 47.19% HHRDE, 5.19% CS and 16.45% STEM.
- HarvardX and MITx (N=565): 26.55% GHSS, 27.07% HHRDE, 7.8% CS
260 and 38.58% STEM.

3.4. Analytic Sample

To enable comparison between cohorts across the study we define the following four populations of participants, where participants are defined as learners that registered for and viewed a course at least once:

- 265 • **Edraak Rest:** Participants who connected to Edraak from outside Arab countries. We are interested in this subpopulation because these are presumably Arabic-speaking learners (since most content in Edraak is in Arabic) residing outside Arab countries (N = 12,203).
- **Edraak Arab:** Participants who connected to Edraak from Arab countries
270 (N = 589,817).
- **MITxHx Arab:** Participants who connected to MITx or HarvardX courses from Arab countries (N = 120,868).
- **MITxHx Rest:** Participants who connected to MITx and HarvardX courses from outside Arab countries (N = 3,267,199).

275 3.5. Multi-Platform MOOC Learning Analytics

Much of MOOC research over the past five years has been conducted in studies of single higher education institutions, and even the largest studies have aggregated data from within a single MOOC provider. In this study, we employ multi-platform MOOC learning analytics, a process that we developed and applied in a previous case study [37]. This process leverages commonalities across
280

MOOC learning and content management systems to allow research teams to create common data formats, agree upon analytic methods, and then generate aggregate data, produced through identical processes, that can allow for “apples-to-apples” comparisons between different MOOC platforms. The goal is to, as closely as possible, conduct the same analysis within the two MOOC environments. This is increasingly important in the field of the social sciences, since human factors can add a great variability in the outcomes of each study [18]. We applied the following procedure:

1. Both research teams shape their data into the same format (see Subsection 3.3). For MITx and HarvardX data we use *edx2bigquery* [38] framework to transform the raw edX logs into this person-course format. Edraak uses a simplified and adapted version of this framework to transform the Open edX data from their platform into this same person-course file.
2. We collaboratively generate a common Jupyter notebook that receives as input the common person-course data format.
3. The common script outputs aggregate data from both platforms that is merged together for the joint analysis.
4. Each research team processes survey data independently and shares aggregate results for analysis.
5. We conduct the joint data analysis of the two platforms together.

4. Results

We organize this section by research question. In response to research question 1—asking whether regional MOOC providers help narrow educational and gender gaps for the local population—we found narrower educational and gender gaps in Edraak, who also reached a younger population. We unpack this finding in Subsection 4.1. Research question 2 asked whether regional MOOC providers have higher levels of engagement, participation, and completion for the local population, compared to global providers. We found that Edraak reached a greater proportion of local and Arab learners, and that completion

310 trends were higher than comparable trends in the global provider. We describe
these findings in Subsections 4.2 and 4.3. Research question 3 explored the de-
gree to which course level and individual level factors were related to the access
and engagement indicators. We found that observed differences in completion
and engagement between edX-licensed courses and home-produced courses in
315 Edraak remained after disaggregating by course category. We describe these
differences in Subsections 4.4 and 4.5. We also report on differences in course
completion on the basis of English proficiency in MITxHx and the Edraak en-
gagement survey. We find that Arabic content is an important factor for local
learners but course fit matters too. We summarize these findings in Subsection
320 4.6.

4.1. Regional Provider Edraak Attracts Younger Learners, Females, and Learn- ers with Lower Levels of Educational Attainment Compared to Global Providers

Level of education. In the left panel of Figure 1 we display a boxplot with the
distribution of level of education, organized by cohort. We map the level of edu-
325 cation of learners to ‘Doctorate’, ‘Master’, ‘Bachelor’, ‘Associate’, ‘High School,
Junior High School or Elementary School’ or ‘Other’, and the visualization also
follows that order (i.e. from the highest to the lowest level of education). The
plot shows that Edraak attracts a larger proportion of learners with a lower
level of education compared to those in MITxHx. More specifically, around 13%
330 of the learners of **Edraak Arab** have a Master or Doctorate, compared to the
19.6% of **MITxHx Arab** and the 25.5% of **MITxHx Rest**. Additionally, if we com-
pare these trends to the rest of Edraak learners, the cohort **Edraak Rest** has
a slightly higher percentage of learners with Doctorate and Masters degrees, a
figure similar to the **MITxHx Arab** cohort. However, they also have a higher
335 percentage of learners with lower levels of education.

Gender. In the right panel of Figure 1, we display the gender distribution by
cohort. The visual inspection shows that **Edraak Arab** cohort reaches a higher
proportion of female than male learners (50.67% female and 48.93% male), com-
pared to **MITxHx Arab** (27.85% female and 71.89% male) and **MITxHx Rest**

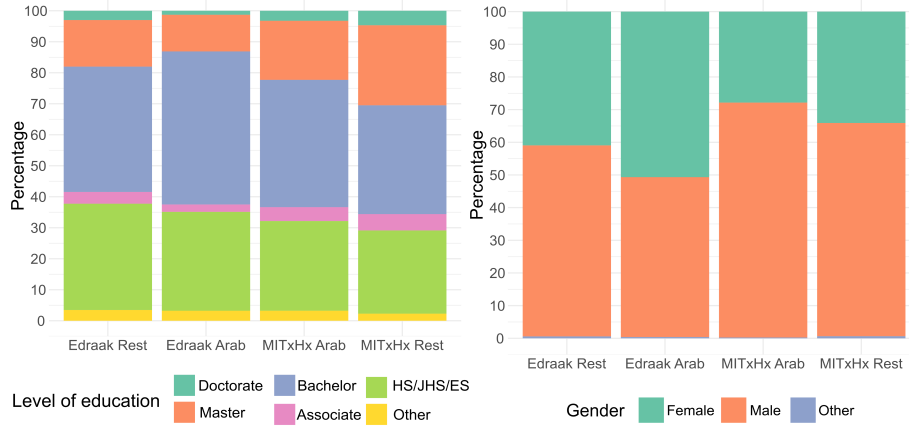


Figure 1: Distribution of gender and level of education by cohort.

340 (34.09% female and 65.26% male). This difference is not as pronounced in Edraak Rest (40.88% female and 58.39% male).

Age. Figure 2 shows on the left a boxplot on the age distribution by cohort, where Edraak Arab has the lowest age (mean = 26.6 and median = 25), followed by MITxHx Arab (mean = 28.1 and median = 26) compared to the considerably higher distribution of MITxHx Rest (mean = 32 and median = 29). Similarly to what we have found with the rest of demographic variables, Edraak Rest shows a different trend with older learners (mean = 29.1 and median = 28) than Edraak Arab. This distribution is transformed into age group intervals (< 18 , $[18, 30)$, $[30, 45)$, $[45, 65)$ and ≥ 65) on the right stacked bar chart of Figure 2. This plot shows that 14.38% of learners are above 45 years for MITxHx Rest compared to 5.51% of MITxHx Arab learners and 3.85% of MITxHx Edraak learners. Moreover, the percentage of learners younger than 30 is 71.13% for Edraak Arab and 66.81% for MITxHx Arab compared to the 52.73% of MITxHx Rest.

350

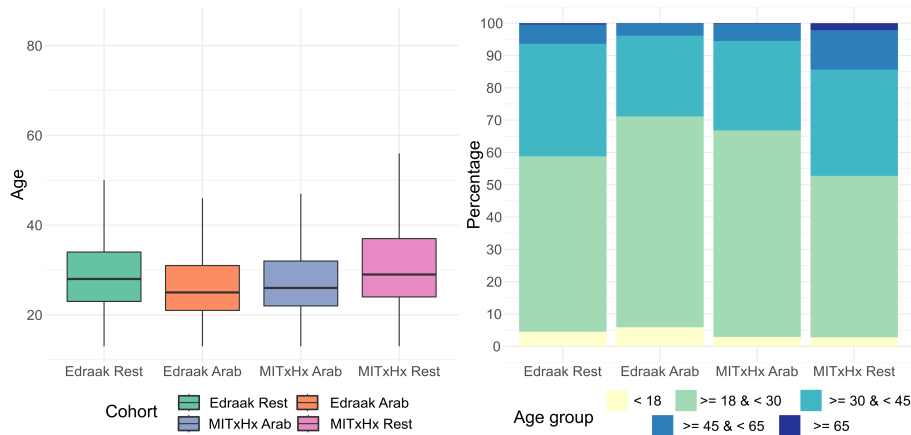


Figure 2: Age distribution.

355 *4.2. Edraak Reaches More Local Learners and Participation is Less Related to
HDI Compared to Global Providers*

We examined differences in participation of the local population in each platform. While MITxHx Arab reaches 120,868 learners in Arab countries which represents a ratio of 285.9 participants per million inhabitants in the region, 360 **Edraak Arab** reaches 589,817 learners residing in Arab countries which represents 1395.3 participants per million inhabitants in the region— nearly five times more than MITxHx Arab.

Table 1 highlights the top and bottom five countries in terms of representation of Arab learners divided by **Edraak Arab** and MITxHx Arab. The representation is highest in Jordan **Edraak Arab** with more than 9,000 participants 365 per million inhabitants. We can observe the contrast by examining the top and bottom countries, which brings up the differences in terms of human development of each country within the Arab region. We find a moderate-high Pearson Correlation of 0.69 ($p < 0.001$) between the HDI index of the country and the 370 number of participants per million inhabitants. The relationship is attenuated in the MITxHx Arab with a value of 0.26 ($p = 0.23$).

Edraak Arab Learners (N = 589,817)			MITxHx Arab Learners (N = 120,868)		
Country	Percentage from total	Participants per million	Country	Percentage from total	Participants per million
Egypt	29.27%	1,737	Egypt	31.52%	383
Jordan	15.27%	9,093	Saudi Arabia	12.31%	443
Saudi Arabia	13.9%	2,444	United Arab Emirates	12.25%	1,552
Morocco	8.65%	1,410	Morocco	10.59%	354
Algeria	8.51%	1,195	Algeria	5.17%	149
Bahrain	0.18%	694	Yemen	0.72%	30
Somalia	0.11%	42	Libya	0.6%	111
Mauritania	0.08%	109	Djibouti	0.25%	315
Comoros	0.02%	119	Mauritania	0.12%	32
Djibouti	0.02%	104	Comoros	0.01%	11

Table 1: Top five and bottom five countries in representation of Arab participants by cohort.

4.3. Edraak Participants Demonstrate Higher Exploration and Completion Rates Compared to Their Arab Peers in edX

Table 2 shows the percentage of viewed, explored and completed per cohort. Comparing completion rates between edX and Edraak suggests higher exploration and completion among the **Edraak Arab** cohort where the percentage of courses completed is triple that of **MITxHx Arab** and double that of **MITxHx Rest**. We see a similar trend among the percentage explored, though the relationship is slightly attenuated. The percentage of learners that viewed the course is higher in the **MITxHx** cohorts, which indicates a higher proportion of students at Edraak signed up for courses but did not access the course content. Similar to previous correlational results, we find a moderate correlation of 0.55 ($p = 0.007$) by country within **MITxHx Arab** for the percentage of completed courses and the HDI. The correlation is attenuated in **Edraak Arab** with a value of 0.45 ($p = 0.03$).

Cohort	Percentage viewed	Percentage explored	Percentage completed	Participants per million inhabitants
Edraak Rest	43.87%	33.39%	10.01%	1.69
Edraak Arab	39.12%	34.72%	10.66%	1395.3
MITxHx Arab	53.65%	15.94%	3.15%	285.93
MITxHx Rest	57.61%	19.43%	4.85%	453.14

Table 2: Participation and completion by cohort.

4.4. Differences in Engagement Persist at the Course Level

We use the high level course categorization reported in Subsection 3.3 [4] to examine participation and completion. In Table 3, we see similar completion trends among Edraak Arab and MITxHx Arab, and lower completion among MITxHx Arab compared to MITxHx Rest. Overall, we find lower completion in CS and STEM courses compared to that of GHSS and HHRDE. MITxHx Arab learners had lower completion rates in CS and STEM compared to their peers in Edraak Arab, while completion of STEM courses is lower for Edraak Arab compared to MITxHx Rest.

We also examined learners’ interests using registrations per course. Registration is higher in GHSS and HHRDE courses in Edraak Arab, compared to MITxHx Arab where there was higher registration in CS and STEM courses, whereas in MITxHx Arab we see a peak of interest for CS courses and similar interest for STEM, GHSS and HHRDE courses.

4.5. Higher Engagement in Edraak-Produced Courses Compared to Local edX Licensed Courses

We capitalize on the fact that Edraak was offering eight courses that were licensed directly from edX to compare participation and completion with other courses directly produced by Edraak. These licensed courses were ‘locally-adapted’ as much as possible by translating all textual content in the courses to Arabic and by editing images and visualizations to change any English text to Arabic. In response to the challenge of Arabizing the visual and audio content

Topic	Cohort	Registrations per course	Percentage viewed	Percentage explored	Percentage completed
CS	Edraak Rest	197	46.59%	32.3%	7.19%
	Edraak Arab	9739	41.68%	25.62%	6.92%
	MITxHx Arab	2964	57.27%	9.91%	1.21%
	MITxHx Rest	70203	61.41%	11.78%	2.75%
STEM	Edraak Rest	208	42.97%	19.28%	3.24%
	Edraak Arab	10674	37.45%	18.75%	3.61%
	MITxHx Arab	706	54.37%	18.55%	2.31%
	MITxHx Rest	14324	58.39%	23.75%	4.16%
GHSS	Edraak Rest	349	46.84%	38.09%	11.8%
	Edraak Arab	15500	43.53%	40.09%	13.38%
	MITxHx Arab	665	49.63%	19.53%	5.62%
	MITxHx Rest	15209	55.78%	22.6%	6.42%
HHRDE	Edraak Rest	328	41.73%	33.16%	10.47%
	Edraak Arab	18126	36.83%	35.08%	10.55%
	MITxHx Arab	665	49.63%	17.42%	5.84%
	MITxHx Rest	15201	53.3%	21.61%	7.39%

Table 3: Participation and completion statistics by course category and cohort.

Licensed from edX?	Enrolments per course	Participants per course	Percentage viewed	Percentage explored	Percentage completed
No	16363	6440	39.36%	35.28%	10.94%
Yes	15009	4824	32.14%	9.25%	1.03%

Table 4: Participation and completion comparison of courses produced by Edraak (N = 223) from those courses licensed from edX (N = 8).

of videos, Edraak provided Arabic transcripts that learners can read alongside the videos. Table 4 compares learner activity in licensed versus locally produced courses. The data show 11 times more completion in courses produced by Edraak compared to those licensed from edX, around three times more exploration and greater overall enrolments per course.

4.6. English Proficiency, Course Participation, Course Completion, and Learner Preferences

This subsection explores trends in course participation and completion on the basis of English proficiency, and self-reported preferences on Edraak’s learners. The first row of Table 5 shows the average values of self-reported English proficiency by region based on the MITxHx survey (from 0 to 4), where the Arab population has one of the lowest (3.11) together with Latin America (2.98). Additionally, Figure 3 shows the percentage of explored, completed and certified in MITxHx by self-reported English proficiency level (encoded using color) and divided by the two cohort (encoded by linetype). Overall, we can see that in both MITxHx Rest and MITxHx Arab cohorts there is a correlation between the English proficiency and funnel of participation. However, this relationship appears to be larger for MITxHx Arab cohort. For MITxHx Rest cohort (denoted by solid lines), the level of completion is similar for learners reporting an intermediate, proficient and fluent level of English proficiency. However, among MITxHx Arab cohort (denoted by dashed lines), there is a gap between the weak, basic, intermediate and proficient levels of proficiency, while for Arab learners reporting fluent English proficiency, the completion levels that are similar to

Survey question	Region						
	<i>Africa</i>	<i>Arab</i>	<i>Asia</i>	<i>Europe</i>	<i>Latin America</i>	<i>Northern America</i>	<i>Oceania</i>
English proficiency	3.65	3.11	3.25	3.38	2.98	3.88	3.85
Threat country	0.52	0.68	0.65	0.38	0.71	0.26	0.23

Table 5: Average value of survey responses per region.

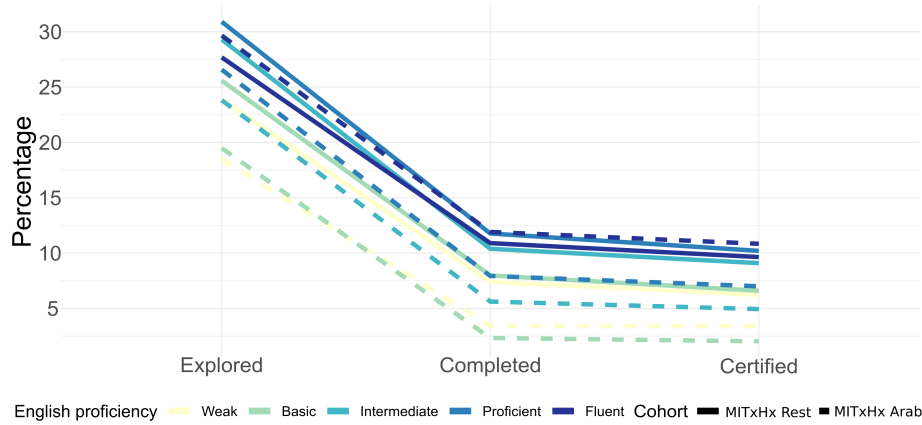


Figure 3: Analysis of completion based on English proficiency in MITx and HarvardX. Line-type codifies cohort.

MITxHx Rest.

Based on the survey responses from Edraak, 99% of learners speak Arabic fluently or have native proficiency, compared to 19% reporting speaking English fluently, 52% moderately and 23% as a beginner. Given the large proportion of Arabic speakers among Edraak learners, we were interested in examining whether language of instruction was related to learners' MOOC provider preferences. The proportion of learners who registered on each platform are shown in Figure 4. We see that 56% of learners who registered on Edraak have not registered on other platforms. For Edraak learners who reported registering in another platform, Arabic MOOC platform, Rwaq with 26% has the highest registration rate among respondents, and 32% of the respondents indicated that they have registered on at least one of the global English-based platforms.

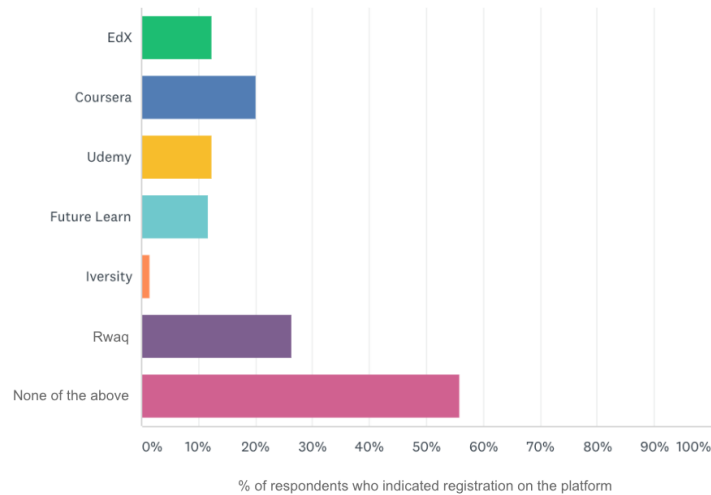


Figure 4: Percentage of Edraak’s learners that have registered in each platform.

Registering on a global MOOC platform is positively correlated with English proficiency ($0.21, p < 0.001$) as reported in the survey.

445 In Figure 5, we present Edraak learners’ self-reported top reasons for registering on Edraak. Edraak learners cite the most common reason for registering on Edraak is “The courses are relevant to my needs” (44%) followed by “Content is in Arabic” (37%). Other common reasons were the “Diversity of the offered courses” and the “Universities that sponsor the courses” with 37% and
450 32% respectively.

In Figure 6, we depict Edraak learners’ perceptions of the merits of different MOOC providers. We normalize these percentages based on the number of learners that reported registering on each platform, e.g. based on the number of learners that registered for edX, what percentage of them think that edX
455 is the platform that fulfills each of the criteria. This weighting allows us to more fairly compare learner preferences in the different providers in and not weight preferences towards Edraak (since all survey participants registered in Edraak). It should be noted that many Edraak participants did not register for an English-speaking provider. The survey results show that Edraak learners

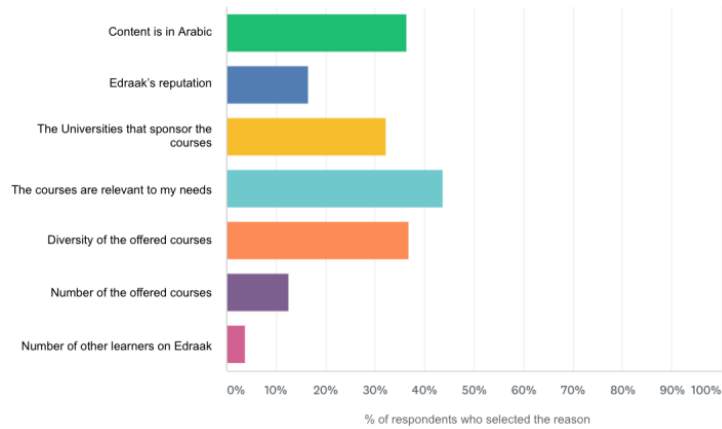


Figure 5: Percentage of respondents that selected each item as one of the most important reasons to register in Edraak.

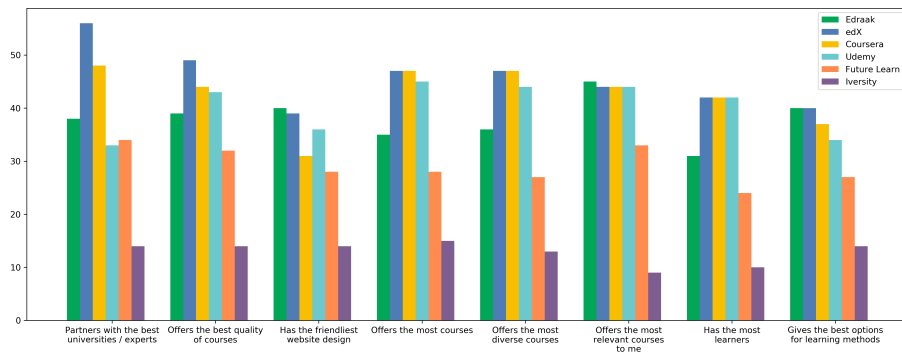


Figure 6: Percentage of learners that voted the best platform for each question item.

460 value English-language platforms for their expertise and selection of courses, particularly edX and Coursera, but also Udemy to some extent, reporting that these providers “Partners with the best universities/experts”, “Has the most learners” and “Offers the best quality”, “the most diverse” and “the most number of courses”. Edraak learners valued Edraak for its accessibility and relevance
 465 selecting Edraak as the provider that “Has the friendliest website design”, “Offers the most relevant courses for me” and “Gives the best options for learning methods”.

5. Discussion

We found that **Edraak** is reaching less educated, younger, and more female
470 learners in the Arab world compared to the MITxHx platform. We found that
Edraak Rest had a higher number of highly educated learners, which might
be due to highly educated Arab learners residing outside Arab states. Edraak
had a high presence among learners per million inhabitants in Jordan, which
475 and marketing initiatives. The finding that more female learners are accessing
MOOCs may signal the promise of MOOCs as a workforce development tool,
particularly given the high unemployment among women in the Arab world
[39]. The relationship between MOOC demand and unemployment has been
documented in previous work [40] and could be one factor linked to higher female
480 enrolment. If in fact female participation in MOOCs is linked to workforce
development, this could lead to numerous positive outcomes in the development
of the region [41].

While we documented higher rates of completion for learners in Edraak
compared to MITxHx, both in the aggregate and by course type, we did not
485 conduct an extensive qualitative study of course offerings to address important
variation in instructional pedagogy, length or difficulty and thus our findings
regarding completion are limited. It is possible that differences in course diffi-
culty are related to overall higher completion rates among Edraak participants.
For instance, Edraak offers fewer CS and STEM courses (where completion is
490 lower) compared to GHSS and HHRDE courses (where completion is higher).
This observed difference in course offerings across category may also be linked
to greater female participation since since prior studies have identified lower
female enrolment in STEM and CS courses [4].

Edraak participants also showed different levels of engagement with and
495 preference for locally-produced courses compared to edX-licensed courses. This
finding underscores the importance of developing localized educational content.
The major difficulty that Edraak learners reported in edX-licensed courses was

the cognitive load of watching courses with subtitles, since they had to split their attention between the video and transcript. Therefore, it seems that sub-
500 titling English content in the local language might still represent an important barrier for some learners. Educational content that does not require students to synchronize multiple separate sources of information would facilitate learning.

We also found that the well-documented relationship between MOOC participation (learners per million inhabitants in a country) and HDI (i.e. the more
505 affluent the country, the greater the number of MOOC participants) evident in MITxHx Arab, is diminished in the Edraak Arab cohort. However, the relationship between percentage of completed courses and HDI present in MITxHx Arab, was lightly attenuated but still persisted in Edraak Arab. Thus, Edraak helps to break the access barrier but learners still struggle to complete based on
510 the HDI of their country, which might indicate that a more scaffolded learning experience might be required for specific subpopulations coming from low HDI countries, even in localized settings.

There are a number of possible factors underlying our observational findings. We presented survey data in order to contextualize the observed trends. An-
515 other hypothesis that might explain low participation and completion in MITxHx Arab is social identity threat. Under this paradigm, an individual has both a personal self identity, but also a collective identity which embraces additional features of the group to which they belong [42]. Social identity threat has been linked to educational attainment gaps in different settings and contexts, such as
520 for first generation college students [43] or racial groups like African American adolescents [44]. The second row of Table 5 shows results of the survey question addressing social identity threat based on the home country, where we present the average value by region which ranges from 0 (no social identity threat) to 4. The results show that the Arab population has one of the largest values of
525 social identity threat (0.68) with only Latin American learners reporting higher levels of social identity threat (0.71). These figures are triple the value of that of their peers in regions with the lowest levels of social identity threat like Northern America (0.26) or Oceania (0.23). Previous work has found that existing

educational attainment gaps in MOOCs based on social identity threat can be
530 reduced with simple activities such as a social-belonging or affirmative value
interventions [5].

Language of instruction may play a role in the observed learning outcomes.
Most Arab learners in the sample speak English as a second language and the
Arab world countries rank in the “very low” category on the EF English Pro-
535 ficiency Index, a global measures of English proficiency [45]. Previous work on
Arab learners have also found self-perceived English proficiency to be signifi-
cantly correlated with foreign language classroom anxiety [46]. As part of the
results, we presented that the gap in the participation funnel based on English
proficiency in MITxHx was larger for Arab learners than for learners from other
540 language groups, and only 19% of Edraak’s learners can speak English fluently.

Despite the English-related barriers that we report, the fact that Edraak’s
content is in Arabic, was not the top-reported reason for Edraak learners’ moti-
vation for enrolling in the regional platform (although it was the second). The
top rated reason for enrolling in regional providers was that the courses being
545 offered in Edraak were relevant to their needs. As mentioned previously, the
match between learner needs and course offerings could be a result of frequent
polling of learners in the region as part of Edraak’s marketing strategy. Edraak’s
learners valued English-speaking platforms for their partnerships with the best
universities, having the most learners and offering the best quality, most diverse
550 and highest number of courses. However, Edraak learners rated Edraak as the
top provider in terms of friendliest website design, the most relevant courses and
best learning methods. Taken together, we hypothesize that regional learners
are responding positively to the localizing of the learning environment, course
content and selection of courses which are targeted to the needs and preferences
555 of the regional population. Future research should further explore the extent to
which the localized learning environment and content affects learning outcomes
like participation and completion among regional learners. The differences in
enrollment based on course category across the regional and global platform
with registration higher in GHSS and HHRDE in Edraak Arab and higher in-

560 terest in CS and STEM courses among MITxHx Arab could also be speaking to learner preferences.

While research and popular discussion has generally focused on the English-language, global providers, our study provides compelling evidence that more attention should be paid to the regional, non-English providers and their efforts to expand online learning among cultural and language groups. Regional providers may be uniquely positioned to expand access to diverse learners.

6. Conclusions

We present a case study employing multi-platform learning analytics [37]—a methodology to execute a common research procedure that allows us to compare MOOC data across the regional provider, Edraak and two global edX providers, MITx and HarvardX.

The goal of the study was to analyze and discuss the demographics and engagement of learners from the Arab states in these two platforms. While numerous research studies (e.g. [47, 21, 5]) have pointed out that that MOOCs did not fulfill their mission to democratize education and reach learners in need, our results are suggestive of the promise of local MOOC providers in achieving this original goal of providing broadly accessible learning experiences for diverse learners, particularly for those learners with limited access to traditional higher education. Edraak reached more female learners, learners with lower levels of education, and those from less developed countries. At the same time, we found that Edraak learners had higher rates of course persistence and completion, compared to the Arab learners in MITx and HarvardX. Additionally, Edraak reached five times more Arab learners compared to global providers and the learners reported that one of the main reasons to enrol in Edraak is that courses match their needs and that the content is in Arabic. Multiple factors from marketing to course design to course rigor to the disciplinary makeup of course offerings to instructional quality could be responsible for some of these differences, even though we raise the idea that regional MOOC providers might

be helpful to start reaching a more diverse, less educated, more female learners
590 among local populations and are best suited to target their instructional needs.

These findings suggest that regional MOOC providers may be understudied and play a very different role in the MOOC ecosystem than global providers. Researchers might consider devoting more resources, time and energy to engage local providers in research studies; particularly because local MOOC providers
595 may not have the resources to independently conduct such studies. One reciprocal question that we pose to the broader MOOC ecosystem is what strategies can global MOOC providers learn from regional MOOC providers to increase equity in the global classroom and design their courses more inclusively. Conversely, what lessons can regional MOOC providers learn from global MOOC
600 providers. Hence, one of the aims of this study is to motivate cross-provider studies and partnerships among more research groups and institutions to shed light on some of these macro learning analytics trends.

We propose multi-platform learning analytics as the first step in pursuing this research agenda. We have shared one approach that bypasses some of the
605 most common challenges of multi-platform and cross-institutional learning analytics research, such as data sharing, privacy and data compatibility. We have demonstrated the insight that can be gained from multi-platform analytics. Additionally, the *cost* required to conduct MOOC research is extensive, when that *cost* is shared by several researchers it lowers the cost of conducting research.
610 Finally, as a result of executing this procedure, we have developed a parallel research process which resulted in an “apples-to-apples” comparison between the trends on each platform which might not have been feasible in two independent research studies.

We are also developing a consortium of researchers with access to data from
615 diverse MOOC providers from across the globe to engage in a multi-platform project with an expanded list of several global and regional partners. The goal of this larger project is to identify global, longitudinal trends in MOOC usage and develop a deeper understanding of more universal and more context-dependent trends and to effectively compare MOOC learning in global and regional con-

620 texts.

References

- [1] G. Siemens, Learning analytics: envisioning a research discipline and a domain of practice, in: Proceedings of the 2nd international conference on learning analytics and knowledge, ACM, 2012, pp. 4–8.
- 625 [2] L. Yuan, S. Powell, MOOCs and open education: Implications for higher education (2013).
URL <http://publications.cetis.ac.uk/2013/667>
- [3] T. R. Dillahunt, B. Z. Wang, S. Teasley, Democratizing higher education: Exploring mooc use among those who cannot afford a formal education,
630 The International Review of Research in Open and Distributed Learning 15 (5).
- [4] I. Chuang, A. Ho, HarvardX and MITx: Four years of open online courses—fall 2012–summer 2016.
- [5] R. F. Kizilcec, A. J. Saltarelli, J. Reich, G. L. Cohen, Closing global achievement gaps in MOOCs, Science 355 (6322) (2017) 251–252.
635
- [6] M. Rohs, M. Ganz, MOOCs and the claim of education for all: A disillusion by empirical data, The International Review of Research in Open and Distributed Learning 16 (6).
- [7] L. Shi, A. I. Cristea, Demographic indicators influencing learning activities in MOOCs: learning analytics of FutureLearn courses, in: 27th International Conference on Information Systems Development (ISD2018), Springer, 2018.
640
- [8] C. McLoughlin, R. Oliver, Designing learning environments for cultural inclusivity: A case study of indigenous online learning at tertiary level,
645 Australasian Journal of Educational Technology 16 (1).

- [9] X. Liu, S. Liu, S. Lee, R. J. Magjuka, Cultural differences in online learning: International student perceptions., *Educational Technology & Society* 13 (3) (2010) 177–188.
- [10] E. E. Grandon, K. Alshare, O. Kwun, Factors influencing student intention to adopt online classes: A cross-cultural study, *Journal of Computing Sciences in Colleges* 20 (4) (2005) 46–56.
- [11] Y.-S. Tsai, P. M. Moreno-Marcos, I. Jivet, M. Scheffel, K. Tammets, K. Kollom, D. Gašević, The SHEILA Framework: Informing Institutional Strategies and Policy Processes of Learning Analytics, *Journal of Learning Analytics* 5 (3) (2018) 5–20.
- [12] S. Buckingham Shum, Unesco policy brief: Learning analytics (no. november), UNESCO Institute for Information Technologies in Education. Retrieved from www.iite.unesco.org/publications/3214711.
- [13] H. Drachler, M. Kalz, The MOOC and learning analytics innovation cycle (MOLAC): a reflective summary of ongoing research and its challenges, *Journal of Computer Assisted Learning* 32 (3) (2016) 281–290.
- [14] L. Breslow, D. E. Pritchard, J. DeBoer, G. S. Stump, A. D. Ho, D. T. Seaton, Studying learning in the worldwide classroom research into edx’s first mooc., *Research & Practice in Assessment* 8 (2013) 13–25.
- [15] A. I. Cristea, A. Alamri, M. Kayama, C. Stewart, M. Alshehri, L. Shi, Earliest predictor of dropout in moocs: a longitudinal study of futurelearn courses, in: *27th International Conference on Information Systems Development (ISD2018)*, Springer, 2018.
- [16] G. Veletsianos, P. Shepherdson, A systematic analysis and synthesis of the empirical mooc literature published in 2013–2015, *The International Review of Research in Open and Distributed Learning* 17 (2).
- [17] T. van der Zee, J. Reich, Open education science, *AERA Open* 4 (3) (2018) 2332858418787466.

- [18] Y. Amir, I. Sharon, et al., Replication research: A “must” for the scientific advancement of psychology, *Journal of Social Behavior and Personality* 5 (4) (1990) 51.
- [19] J. Gardner, C. Brooks, J. M. L. Andres, R. Baker, MORF: A Framework for MOOC Predictive Modeling and Replication At Scale, arXiv preprint arXiv:1801.05236.
- [20] C. Alario-Hoyos, M. Pérez-Sanagustín, C. Delgado-Kloos, M. Muñoz-Organero, A. Rodríguez-de-las Heras, et al., Analysing the impact of built-in and external social tools in a mooc on educational technologies, in: *European Conference on Technology Enhanced Learning*, Springer, 2013, pp. 5–18.
- [21] J. D. Hansen, J. Reich, Democratizing education? examining access and usage patterns in massive open online courses, *Science* 350 (6265) (2015) 1245–1248.
- [22] Authors blinded for review, Title blinded for review, in: *Learning With MOOCS (LWMOOCS)*, IEEE, 2018, pp. 47–50.
- [23] L. Song, J. R. Hill, A conceptual model for understanding self-directed learning in online environments, *Journal of Interactive Online Learning* 6 (1) (2007) 27–42.
- [24] B. Moser-Mercer, MOOCs in fragile contexts, *European MOOCs Stakeholders Summit*.
- [25] K. Abouchedid, G. M. Eid, E-learning challenges in the arab world: Revelations from a case study profile, *Quality Assurance in education* 12 (1) (2004) 15–27.
- [26] G. Cook, R. Linqanti, M. Chinen, H. Jung, National Evaluation of Title III Implementation Supplemental Report: Exploring Approaches to Setting English Language Proficiency Performance Criteria and Monitoring English

Learner Progress, Office of Planning, Evaluation and Policy Development,
US Department of Education.

- [27] D. August, T. Shanahan, Report of the national literacy panel on language
minority youth and children (2006).
- 705 [28] A. Hunt, S. Tickner, Cultural dimensions of learning in online teacher ed-
ucation courses, *Journal of Open, Flexible, and Distance Learning* 19 (2)
(2015) 25–47.
- [29] A. Ogan, E. Walker, R. Baker, M. M. T. Rodrigo, J. C. Soriano, M. J.
Castro, Towards understanding how to assess help-seeking behavior across
710 cultures, *International Journal of Artificial Intelligence in Education* 25 (2)
(2015) 229–248.
- [30] M. M. T. Rodrigo, R. Baker, L. Rossi, et al., Student off-task behavior in
computer-based learning in the Philippines: comparison to prior research
in the USA, *Teachers College Record* 115 (10) (2013) 1–27.
- 715 [31] K.-J. Kim, C. J. Bonk, Cross-cultural comparisons of online collaboration,
Journal of Computer-Mediated Communication 8 (1) (2002) JCMC814.
- [32] Z. Liu, R. Brown, C. Lynch, T. Barnes, R. S. J. de Baker, Y. Bergner, D. S.
McNamara, Mooc learner behaviors by country and culture; an exploratory
analysis, in: EDM, 2016, pp. 127–134.
- 720 [33] G. Hofstede, Dimensionalizing cultures: The hofstede model in context,
Online readings in psychology and culture 2 (1) (2011) 8.
- [34] P. Parrish, J. Linder-VanBerschot, Cultural dimensions of learning: Ad-
dressing the challenges of multicultural instruction, *The International Re-
view of Research in Open and Distributed Learning* 11 (2) (2010) 1–19.
- 725 [35] E. T. Hall, M. R. Hall, Understanding cultural differences:[Germans, French
and Americans], Vol. 9, Intercultural press Yarmouth, ME, 1990.

- [36] United Nations, Human Development Index (HDI) (2018).
URL <http://hdr.undp.org/en/content/human-development-index-hdi>
- 730 [37] Authors blinded for review, Title blinded for review, in: Proceedings of the Sixth ACM Conference on Learning@Scale, ACM, 2019.
- [38] G. Lopez, D. T. Seaton, A. Ang, D. Tingley, I. Chuang, Google bigquery for education: Framework for parsing and analyzing edx mooc data, in: Proceedings of the Fourth (2017) ACM Conference on Learning@ Scale, ACM, 2017, pp. 181–184.
- 735 [39] World Employment Social Outlook: Trends for Women, Tech. rep., International Labour Organization, Geneva (2018).
URL https://www.ilo.org/global/research/global-reports/weso/trends-for-women2018/WCMS_{_}619577
- 740 [40] T. Tong, H. Li, Demand for MOOC—An Application of Big Data, China Economic Review 51 (2018) 194–207.
- [41] Arab Human Development Report, Tech. rep., United Nations Development Programme, New York (2016).
URL <http://arab-hdr.org/>
- 745 [42] M. B. Brewer, Social identity, distinctiveness, and in-group homogeneity, Social cognition 11 (1) (1993) 150–164.
- [43] N. M. Stephens, M. G. Hamedani, M. Destin, Closing the social-class achievement gap: A difference-education intervention improves first-generation students’ academic performance and all students’ college transition, Psychological science 25 (4) (2014) 943–953.
- 750 [44] T. M. Chavous, D. H. Bernat, K. Schmeelk-Cone, C. H. Caldwell, L. Kohn-Wood, M. A. Zimmerman, Racial identity and academic attainment among african american adolescents, Child development 74 (4) (2003) 1076–1090.

[45] Wikipedia, Ef english proficiency index (2019).

755 URL https://en.wikipedia.org/wiki/EF_English_Proficiency_Index

[46] J.-M. Dewaele, T. M. Al-Saraj, Foreign language classroom anxiety of arab learners of english: The effect of personality, linguistic and sociobiographical variables., *Studies in Second Language Learning and Teaching* 5 (2) 760 (2015) 205–228.

[47] J. Reich, J. A. Ruipérez-Valiente, The MOOC Pivot, *Science* 363 (6423) (2019) 130–131.

Appendix A Survey description

We use the pre-course surveys that are administered as part of all MITx and 765 HarvardX courses at the beginning of the course. We utilize the following two questions:

1. “How worried, if at all, are you that some people in this course will draw conclusions about you, based on what they think of your home country?”
– Single response (Extremely worried | Very worried | Moderately worried 770 | Slightly worried | Not at all worried).
2. “How would you describe your English language skills on the following 5-point scale?” – Single response (Fluent | Proficient | Intermediate | Basic | Weak).

We also incorporated data from a perception survey administered by Edraak 775 in March, 2017. The survey consisted of 18 questions that measure participants’ familiarity with Edraak and other platforms, their perceptions of these platforms’ qualities, and their course enrollment and completion behaviors on them. They survey was distributed to learners who have registered on the Edraak platform within the last 12 months from the survey distribution. The 780 survey is originally in Arabic but questions have been translated into English. We include the following survey items:

1. “Did you previously register on any of the following platforms?” – Multiple choice response (edX | Coursera | Udemy | Future Learn | Iversity | Rwaq)
- 785 2. “Which of the following languages do you speak proficiently? Participants rated their proficiency in Arabic, English, and French on the following 5-point scale:” Single response (Mother tongue | Yes, fluently | Yes, with some difficulty | I can only speak a few words or sentences | I don’t speak it at all).
- 790 3. “Thinking back to the when you registered on Edraak, which of the following reasons for registration were most important to you? Select up to 3 reasons.” – Multiple choice response (Content is in Arabic | Edraaks reputation | The Universities that sponsor the courses | The courses are relevant to my needs | Diversity of the offered courses | Number of the offered courses | Number of other learners on Edraak | Ease of using the platform).
- 795 4. “For each of the following statements, please select the platform(s) to which you feel the statement applies.” – Multiple choice response (Partners with the best universities | experts | Offers the best quality of courses | Has the best interface design | Offers the most courses | Offers the most diverse courses | Offers the most relevant courses to me | Has the most learners | Gives the best options for learning methods).
- 800