Learning Socially

Q8.1 What did distance learning look like before the Internet, and what are some examples?

The *postal mail system* was the earliest technology used for distance learning. The first known example was in the 1840s – roughly a century and a half before the popularization of the Internet – when the Englishman Sir Isaac Pitman used postal mail to teach students his system of phonetic shorthand. Pitman would mail his students postcards with messages written on them, and their task would be to transcribe the text to/from shorthand and mail it back to him, so that he could correct their submissions and send them back again.

The concept of *feedback* was crucial to Pitman's system, mimicking an important characteristic of any classroom: addressing student misunderstanding during the course. The early versions of institutionalized distance learning lacked this feature for the most part, though. Colleges would mail students the course materials and have them report to testing centers to take the exams, with little to no contact between instructor and student (or student and student). A few universities in England had degree programs operating like this by the 1850s, and the model spread to the US by the 1890s. The most notable example is perhaps the International Correspondence School (ICS), founded in 1888 as a school for coal mining. ICS is still around today, providing vocational training to tens of millions of students worldwide. But these students no longer receive their materials through the postal mail; instead, they access them over the Internet, exemplifying how distance learning has evolved with technological innovation.

The second technology that was used to facilitate learning was *radio*. Instructors and/or educational institutions would broadcast audio lectures over the airwaves. In order to do so in the US, an institution first has to get a license from the FCC that specifies the frequencies it can operate in, like how the cellular providers that we looked at in Chapter 1 need to get permission to operate phones in certain ranges. In 1921, the FCC granted the first educational radio license to Latter Day Saints' University, and between 1920 and 1945, over 200 such licenses would be granted to colleges, universities, and school boards.

Despite their general popularity at the time, by 1940, there was only one *credit-bearing* college course offered by radio in the US. Reluc-

tance to offer certification to students learning through this medium abounded, even if people would report to official testing centers to take exams. In fact, there are many policy debates surrounding certification through distance learning, some persisting to the present day. One of the biggest is the issue of *authentication*, *e.g.*, making sure people are doing their own work.

By 1950, the educational radio movement was dying out. At this time, a third learning technology came (quite literally) into the picture: *television*. With the advent of video broadcasting, the visual portions of a lecture (*e.g.*, writing on the board) could be transmitted to students together with the audio. Some universities began offering this for students, the most notorious example probably being the Stanford Instructional Television Network. Formed at Stanford University in 1968, it was first used primarily by students enrolled at the school to *e.g.*, watch a lecture if the student had to miss it.

It wasn't until the 1990s that the most recent and profound vehicle for distance education, the *Internet*, would begin to cause a dramatic change in the landscape of learning. It should come as no surprise that the Internet has become, by far, the most popular technology for distance education, and that it has dramatically increased the enrollment in, and outreach of, these programs.

Q8.2 What were some of the other major events in the evolution of MOOC?

In 2007, two major events unfolded. One was the rise of Khan Academy, started on YouTube by Sal Khan to provide free education to anyone, anywhere. By the count of the number of students who have listened to an instructor's voice, Khan is probably the most heard teacher in history. The other event was the creation of iTunes U by Apple. Even though the number of university courses taken on iTunes U has not been as high as Steve Jobs might have hoped for, it marked another step in the MOOC evolution.

Now, remember the discussion about Stanford recording and broadcasting lecture videos over TV from Q8.1? Well, Stanford Engineering Everywhere started putting these videos online for public consumption, rather than just for the "walled garden," paid-only audience. Then, in summer 2011, three computer science courses at Stanford announced that they would put all their lecture videos online for free,

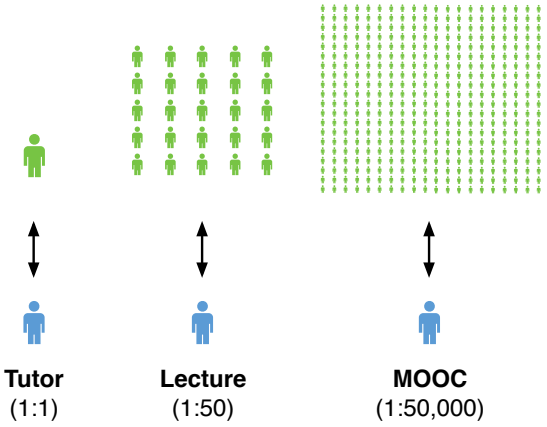

Illustration 26: Three learning scenarios in which the teacher-to-student ratios are vastly different. In tutoring, the instructor (tutor) will teach one or very few students at a time. In a lecture, the instructor (teacher) will typically have more than a few students in a class, say 50. In a MOOC, the number of students increases by orders of magnitude, to typically tens of thousands.

together with multiple choice questions that provide instant feedback at the end of each short video clip, and discussion forums for students (and instructors) to interact. This whole package was offered at no charge and to anyone in the world with an Internet connection. The response was overwhelming, with each course enrollment totaling over 100,000 from across the globe.

Q8.3 Why does the "massive" component of MOOC make differentiation difficult?

The size of the student bodies can be *massive*, orders of magnitude larger than what we are used to in traditional classrooms (see Illustration 26). This means that the **teacher-to-student ratios** in MOOCs are very small, usually just fractions of one percent, *i.e.*, for every one member of the teaching staff in a course, there could be thousands of students.

To get an idea of what the teacher-to-student ratios tend to be in

MOOCs, we took most of the courses offered on Coursera in summer 2013, and compared the number of instructors that participated in class discussions to the number of students that did (we will explain the mechanism for discussions shortly). In doing so, we found that the average ratio for a course was 0.0035, meaning that there were fewer than 4 instructors for every 1000 students on average. And since most students do not even participate in discussions in the first place, the *actual* ratios will be much smaller yet!

Imagine being responsible for thousands of students at a time? It would be virtually impossible to accommodate all the requests that would arise on an individual basis. This would be the case even if it was somehow possible to bring all the students together *in person*, which would at least give the instructor the opportunity to learn about them through face-to-face interaction. But in MOOCs, we have two additional issues as well.

Q8.4 Why does the "online" component of MOOC make differentiation difficult?

Social learning is critical for students. What you may not realize is that it's needed for *teachers*, too. Interacting with students allows a teacher to identify differences that exist between them, which helps to determine if and how to assist on an individual level. Teachers are also trained to recognize differences in how their students learn best. But if a student is confused about something and chooses not to convey it to the teacher, how will the teacher be able to help? She can't, because she won't be aware that a problem exists, at least not until after the test when it's already late in the game. This is why interaction is so important.

Unfortunately, it is much more complicated to get social learning to work effectively in online courses. Why? Well, the interaction is *online*, as opposed to face-to-face, and is *asynchronous*, as opposed to back-and-forth dialogue occurring in real-time. The Internet becomes a medium that physically separates people, as shown in Illustration 27, making it significantly harder to get to know one another. Of course, there are advantages to asynchronous communication too, especially since it gives people time to think before responding, instead of being expected to do so immediately. But in traditional classrooms, we can

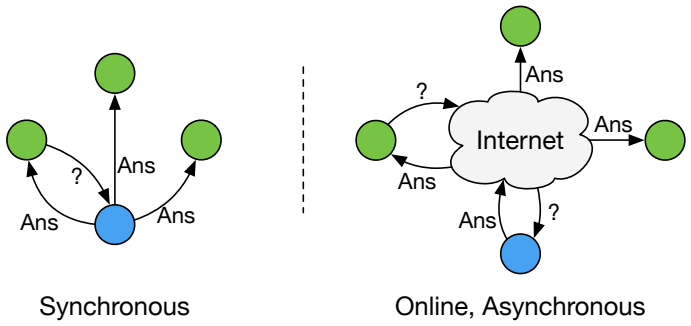

Illustration 27: The difference between synchronous learning in a class environment (left), and online, asynchronous learning (right). In the classroom, when a student asks a question (denoted by the arrow with a ?), the teacher will answer in real time, and all the students can listen. In an online course, the Internet becomes a medium separating humans: when a student asks a question, it is posted on the Internet, and the teacher will retrieve it the next time she logs on.

have the best of both worlds, whereas in MOOCs we are confined to just the online, asynchronous case.

Making matters worse is the fact that some MOOCs are offered *on-demand*. In those cases, students can sign up for the course whenever they want, and proceed through the material on their own time and schedule. While this can be convenient for students, it also introduces asynchrony on a broader timescale: at any given time, students will be different parts of the way through the course. On the other hand, *session-based* online courses usually have a fixed, weekly schedule that everyone is supposed to follow, based on a syllabus provided by the instructor. As with traditional classrooms, the syllabus will specify *e.g.*, which lectures are covered which week, and which assignments are due when.

So, in session-based MOOCs, we can expect that users will at least be focusing on the same material each week, as long as the course has provided incentives for them to stick to the timeline. What would these incentives be? A standard one is in the form of a *certificate of completion*. The criteria for awarding a certificate will vary depending on the instructor's preference, but a student's final grade is typically the determining factor. Sticking to the schedule is necessary to get a

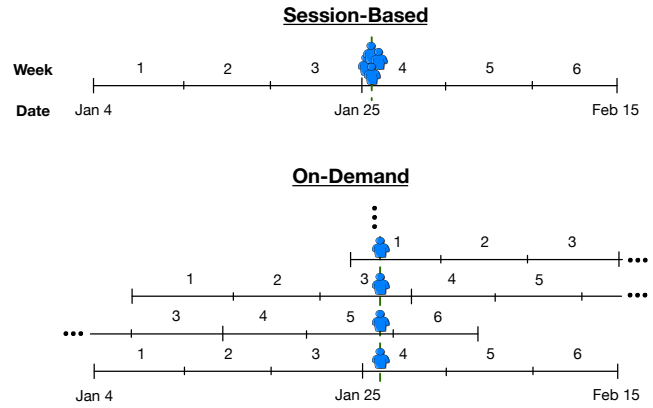

Illustration 28: In a session-based MOOC, the course has fixed start and end dates, and students will be held to that schedule. In an on-demand MOOC, each student can choose her own start date, and the schedule will adjust based on that.

high final grade.

Being the primary channel for social learning, forums are an important part of MOOC. But forum posting is much different than face-to-face dialogue, and lacks the advantage of real-time communication, making it very hard for the instructor to get to know the students. In a session-based MOOC offering a certificate, why can't we just make it a requirement that everyone sign on at a certain time of day, to make discussions occur more synchronously? Think about how difficult this would be to enforce:

- People are in *different timezones*: Whatever time it is, it is always the opposite time of day halfway around the globe. The ideal time for someone in New York on a lunch break (*e.g.*, 1 pm local time) would be very inconvenient for someone in Hong Kong (1 am).
- People have *different responsibilities*: Even within the same timezone, MOOC students may have vastly different lives. What time of day would be convenient for two employees, one who works during the day and the other who does a night shift?

Q8.5 Why does the "open" component of MOOC make differentiation difficult?

Since these courses are *open* to the general public, the student bodies will exhibit diverse demographics, with vastly different geographical and educational backgrounds, and different expectations of what they will get out of the course.

In a traditional classroom, the student bodies are, by design, much more homogeneous. The teacher can expect that each of her students have met some level of background knowledge deemed necessary to understand the material that she will be teaching. This is the premise behind the sequential nature of education we are required to indulge in through adolescence: schools have set curricula that we follow, with the set of classes we take each year for the most part predetermined. In college, our paths will deviate substantially depending on our majors, but each course still has explicitly defined and enforced *prerequisites*. Sometimes this comes in the form of the class number, *e.g.*, in Calculus II it is assumed you have already taken Calculus I. Typically, the transcript of your previous courses and grades can automatically bar you from enrollment if your credentials make you ineligible.

In MOOCs, for the most part, curricula do not exist. Sure, instructors will write in a list of topics they expect that you already know before taking their class. But can we assume that this will be sufficient deterrent to those not meeting the prerequisites? Probably not: if the course is free, there's not much to lose from signing up anyway. But would we even want to bar a student from signing up? Not necessarily: we must keep in mind that students have *different intentions* for taking a MOOC in the first place. Many people are on the standard path of taking every lecture and exam to get a certificate, but others may just be interested in browsing through one or two lectures that piqued their interest in the first place, and do not plan to watch every lecture. Others yet may very well know they do not meet the prerequisites, but are willing to give it a go anyway, with the ambition to pick up *both* the background *and* the current topics along the way.

The end result is that the student demographic in an average MOOC is *much more diverse* than in a traditional class. People will sign up from all over the world, in all different age groups, with varying educational backgrounds, different native tongues, and diverse intentions for why they are there in the first place. With such a broad range of cases to cover, the instructor's job of differentiating learning becomes more

difficult, yet again.

In a resemblance to curricula, some MOOC providers have begun recommending courses that should be taken in sequence. In a *specialization program*, a student will pay money (usually a few hundred dollars) to enroll in a series of MOOCs (usually 3 to 5) on a specific topic, and upon successful completion of the deliverables, will obtain a specialization certificate in the subject area. In these programs, we can expect the student intentions to not be quite as diverse as in a standard MOOC: enrollment requires upfront payment, so student objectives will be more aligned around obtaining the final certificates from the get-go. In progressing through the sequence of courses, subject knowledge among the students will also converge on a more consistent baseline (*e.g.*, in the third course of the program, we can expect students to at least know the material from the first and second).

Q8.6 What is the implication of one-size-fits-all teaching in MOOC?

With only a few instructors to manage so many students, the students will by-and-large receive the same learning experience. They will see the same lecture videos, the same example problems, the same homework solutions, the same emails, and the same answers to the same questions on the forums. Indeed, as class sizes scale up, the instructional styles become largely one-size-fits-all out of necessity, with (proportionately) smaller amounts of differentiated learning on an individual basis. Further, the students have different backgrounds and different intentions for being in a MOOC, which means that for any one way that the course is explained, it will not match the learning needs for many of them.

When someone isn't getting what they need from the course, we would expect them to lose interest, have poor grades, or both. What we see in MOOC is that many students in a course end up dropping out, with only a small fraction of those who are enrolled at the start of the course seeing it to completion. In other words, MOOCs have low **completion rates**. Completion rates are also different for different MOOCs; in fact, they tend to get *lower* as the enrollment rises. For illustration, let's say that course A has 50,000 students enrolled initially and 5,000 finish. What is the completion rate for A? $5,000/50,000 = 0.1$, or 10%. Now, if course B has 200,000 initially and we see 10,000 finish,

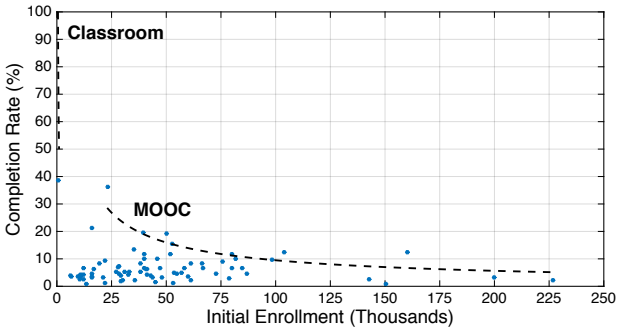

Illustration 29: Completion rate versus initial enrollment observed for 77 MOOCs through June 2015. Here, a student is considered to have completed the course if she received a certificate at the end. We see that the completion rates for MOOCs are rather low, typically below 10%, whereas in a traditional classroom, enrollments are much lower (a few dozen) but completion rates are much higher (towards 100%). The overlaid black curve gives the trend of the highest completion rate for different enrollments, illustrating the scale-efficacy tradeoff of learning.

then what is the rate for B? It is $10,000/200,000 = 0.05$, or 5%. Even though double the number of students completed B than did A, this comes only after *four times* the number of students signed up in the first place.

These numbers are examples, but they are qualitatively characteristic of what has been observed in MOOCs. We show the empirical relationship between enrollment and completion rate for over 75 courses in Ilustration 29, where each of the blue datapoints corresponds to one course. This verifies two key points:

- First is that completion rates in MOOCs are quite low overall, typically less than 10%.
- Second comes from the black curve overlaid on the data, which gives the trend of the highest observed completion for different enrollments. It is decreasing, meaning that for MOOCs with higher enrollment, we generally see lower completion rates. We can compare this too with the case of a traditional classroom, which has orders of magnitude smaller enrollment, but much higher completion.

This second point illustrates what is known as the **scale-efficacy tradeoff** of learning. At the end of the day, while technology has allowed us to have extremely high enrollment in a single course offering, it has (so far) come at the expense of extremely low completion rates. In other words, scaling up has hindered the efficacy of learning.

In making this conclusion, we have assumed that "completion" is a measure of "efficacy." Is this always the case? Not exactly: remember that MOOC students have different intentions for being in the course, not all of them targeting a certificate of completion. The person who enrolls in History of the United States to learn about the American Civil War may only watch the lectures covering 1850 to 1870 and still consider her experience to be perfectly effective. But we'll stick to this assumption for our purposes here, since completion is one of the main ways that traditional classroom outcomes are evaluated.

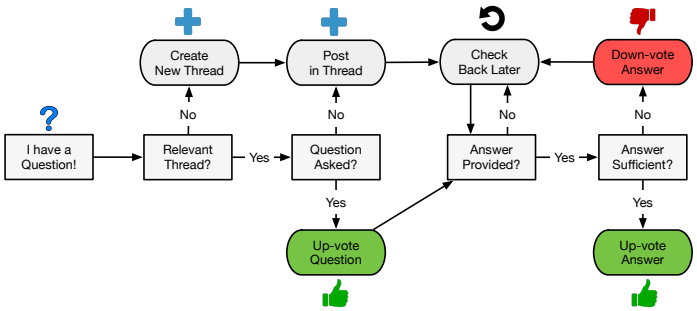

Illustration 30: Flowchart of how a student in a MOOC may navigate the discussion forum when she has a question. The end objective is for the student to receive a satisfactory answer to her question. Until she has one, she will: post the question if it isn't already there, up-vote it if it is there, and up-vote or down-vote answers accordingly, continuing to check back.

Q8.7 Can we put a user's process of navigating a MOOC discussion forum into a flow chart?

In Illustration 30, we give a flowchart of a typical user's forum navigation process. Of course, there are many variations on this that will exist student-to-student, question-to-question, and forum-to-forum, but

our purpose here is just to give an idea of how the process tends to unfold.