Analysis of Repeat Learners in Computer Science MOOCs

CHRISTINE VONDER HAAR, ANA BELL

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

Background

Computational MOOCs are popular, but still suffer from high drop out rates

Two computational courses (6.00.1x and 6.00.2x) offered on EdX have each been offered 10+ times since with thousands of learners per run

We analyzed performance of learners who took the same course multiple times

Research Questions

What are some typical subgroups of repeat learners?

How do repeat learners interact with the course differently?

Is there a difference in performance between different subgroups of repeat learners?

How does an introductory course (6.00.1x) differ from an advanced course (6.00.2x) in repeat learner engagement and performance?

Data Cleaning

Removed all learners who registered but never viewed the course (around 45% of learners)

- We only want to consider *active learners*

Remove *repeat cross references*

- Analyzed this subgroup separately

Removed learners who only completed the graded problems

- Speculated that some learners already knew how to code and were just going through the course for certification

Final data set includes 857,000 learners from 6.00.1x and 122,000 learners from 6.00.2x

Repeat Learner Subgroups

A *repeat learner* took the same course two or more times
- ◦ Around 20% of learners in a course run

A *two-time repeat learner* took the same course exactly two times
- ◦ ~75% of repeat learners are two-time repeat learners

A *repeat cross-referencer* looked at previous runs of a course while taking a current run
- ◦ 2400 in 6.00.1x and 175 in 6.00.2x

A *dual-course learner* took at least one of each course (6.00.1x and 6.00.2x)
- ◦ 62% of learners in 2x took 1x at some point

Repeat learners

- The graphs show the percent of learners who engage in the course in any semester
  - *viewed* (a learner who logs on at least once)
  - *explored* (a learner visits at least half the chapters)
  - *completed* (a learner who earns a passing grade of 55% or more)

- Repeat learners have a higher engagement across the board
- This repeat learner phenomenon affects 1x more than 2x
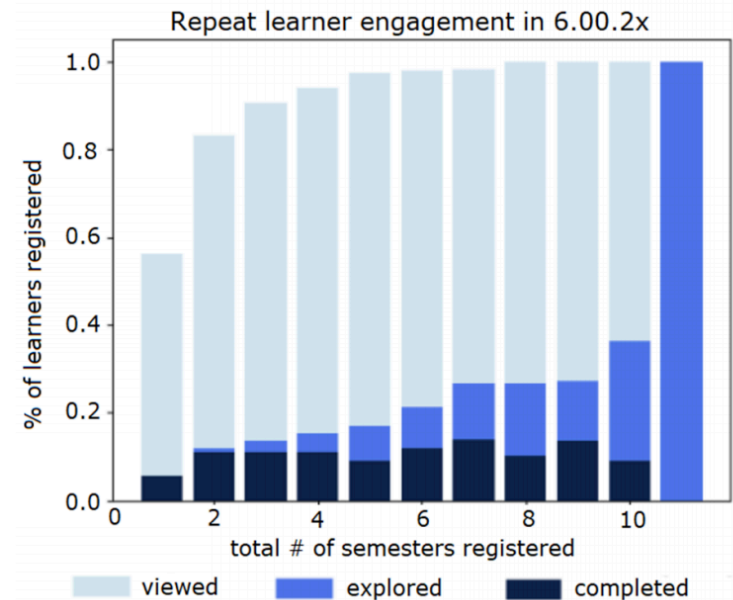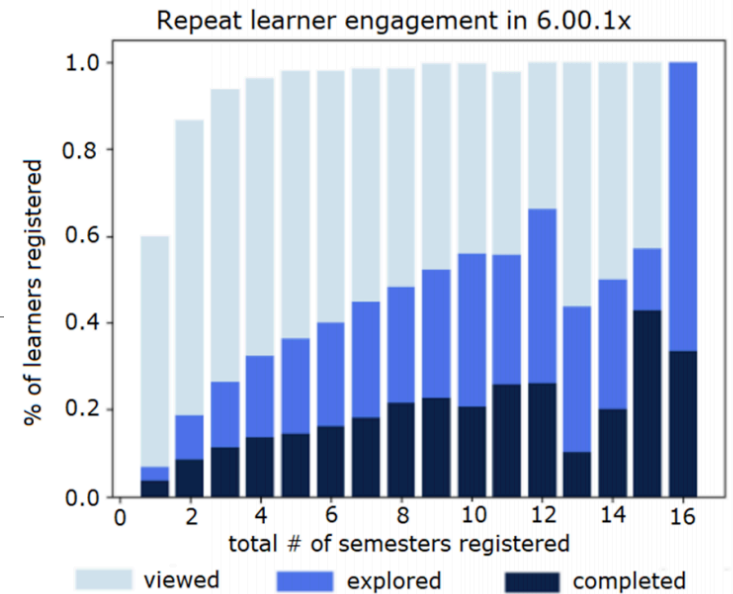
Two-time repeat learners

- The graphs show
  - (top) the number of semesters that occured between a two-time repeat learner's runs
  - (bottom) the completion rate of two-time repeat learners in their second run

- Most two-time repeat learners take the course for a second time right after their intial run
- Learners who take the second run soon after their initial run are generally more successful than those who wait a while between runs.
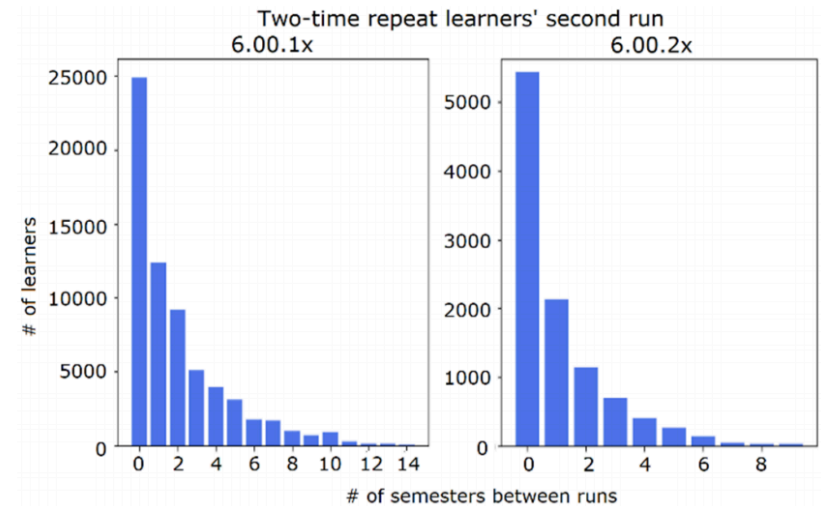

Fig. 2.   The total number of two-time repeat learners vs. how many semesters occurred between their two runs of 6.00.1x (left) and 6.00.2x (right).
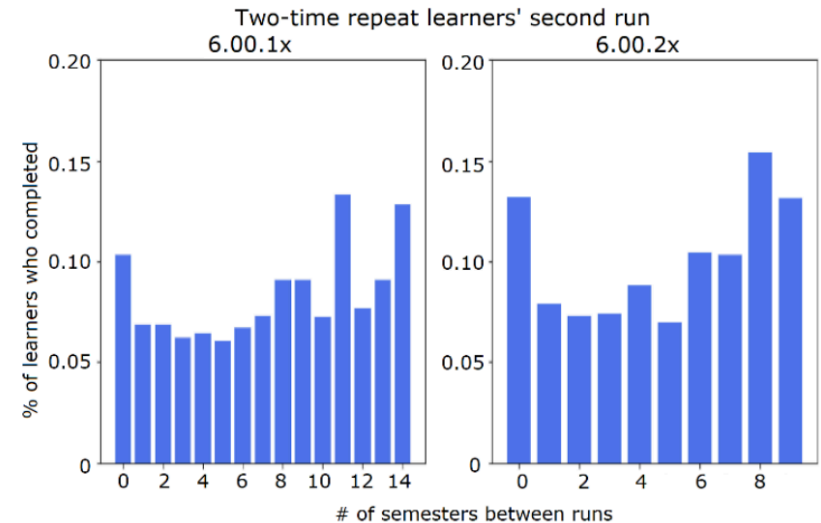

Fig. 3.   The completion rate of two-time repeat learners in their second run vs. how many semesters occurred between their two runs of 6.00.1x (left) and 6.00.2x (right).

Repeat cross-referencers

Repeat cross-referencers eventually complete the course at suspiciously high rates

- repeat cross-referencers eventually complete at a rate of 33% in 1x and 32% in 2x

- average repeat learner eventually complete at a rate of 9.7% in 1x and 11% in 2x

Repeat cross-referencers are often flagged in the run that they complete

- True for 61.4% of cross-referencers in 1x and 72.2% of cross-referencers in 2x

Dual-course learners

TABLE I.   COMPARING LEARNERS WHO TOOK 6.00.1X AND 6.00.2X (DUAL) VS. LEARNERS WHO ONLY TOOK ONE OR THE OTHER (SINGLE).

	6.00.1x Dual	6.00.1x Single	6.00.2x Dual	6.00.2x Single
Mean completion rate	23.5%	3.3%	12.2%	3.1%
Ever completion rate	38.5%	3.8%	15.2%	3.2%
Mean active days	23.5	9.1	14.3	5.7
Mean grade	87.3	81.3	84.6	84.6

- The table shows
  - Mean completion rate: average completion rate per semester
  - Ever completion rate: percent of students who ever take the course and ever complete it
- Dual course learners have higher completion rates than their single-course learner counterparts.

Conclusions

Repeat learners engage with and complete a course more than their single-run counterparts

Two-time repeat learners are more likely to complete the course if they take their second run soon after their first

Repeat cross-references complete the course at suspiciously high rates, indicating possible cheating behavior

Many opportunities for future research
◦ Effects of time- and feature-gating