Engaging with Massive Online Courses

Ashton Anderson, Daniel Huttenlocher, Jon Kleinberg, Jure Leskovec

massive online courses

over 1,000 moocs offered across many platforms 10 million students and counting

massive online courses

potential to revolutionize higher education

...and our understanding of how people learn

massive online courses

but first:

how do students engage with moocs?

can engagement be incentivized?

the data

we studied 6 coursera classes:

3 machine learning (Andrew Ng) 3 probabilistic graphical models (Daphne Koller)



(Thanks to Coursera and the Stanford Lytics Group for sharing the data with us!)

the data

Class	Students	HWs	Quizzes	Lectures	Posts	Start
ML1	64,536	432,052	1,486,566	3,222,074	15,274	4/2012
ML2	60,092	488,554	1,563,301	3,066,189	15,763	8/2012
ML3	112,897	681,569	2,076,354	4,742,864	32,200	4/2013
PGM1	30,385	398,314	794,290	1,564,87	14,572	3/2012
PGM2	34,693	210,199	427,209	1,059,464	7,044	9/2012
PGM3	25,930	172,539	337,657	686,899	4,320	7/2013

Basic course statistics

engaging with massive online courses

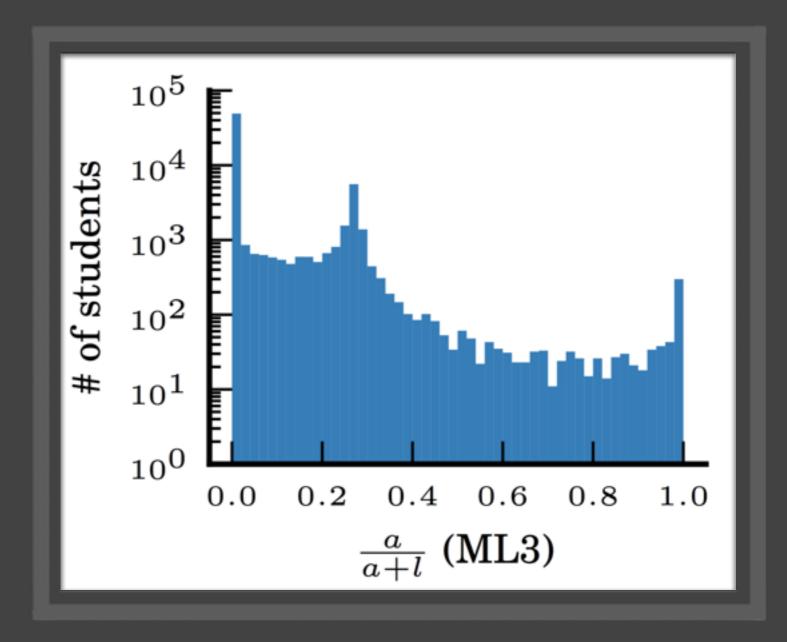
- 1. participation
- 2. performance
- 3. interaction
- 4. intervention

engaging with massive online courses

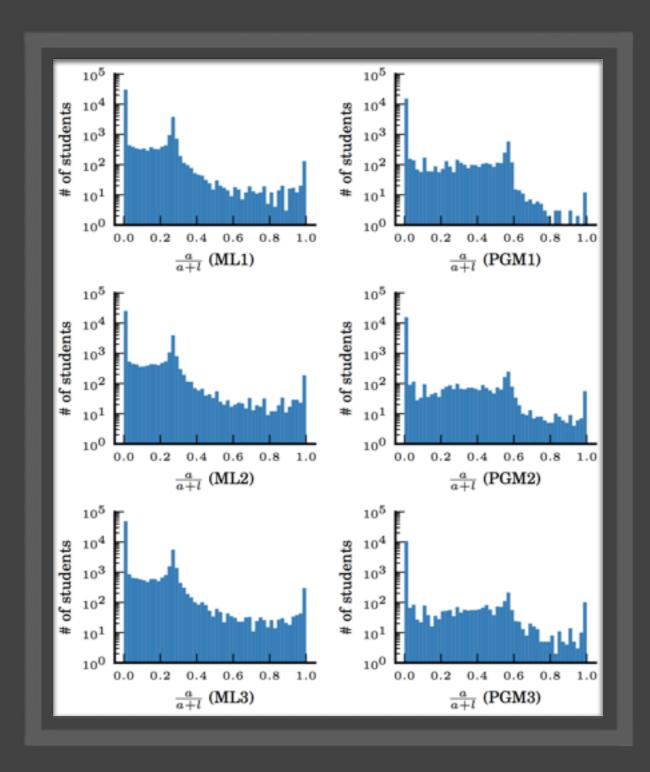
- 1. participation
- 2. performance
- 3. interaction
- 4. intervention

are different students using moocs differently?

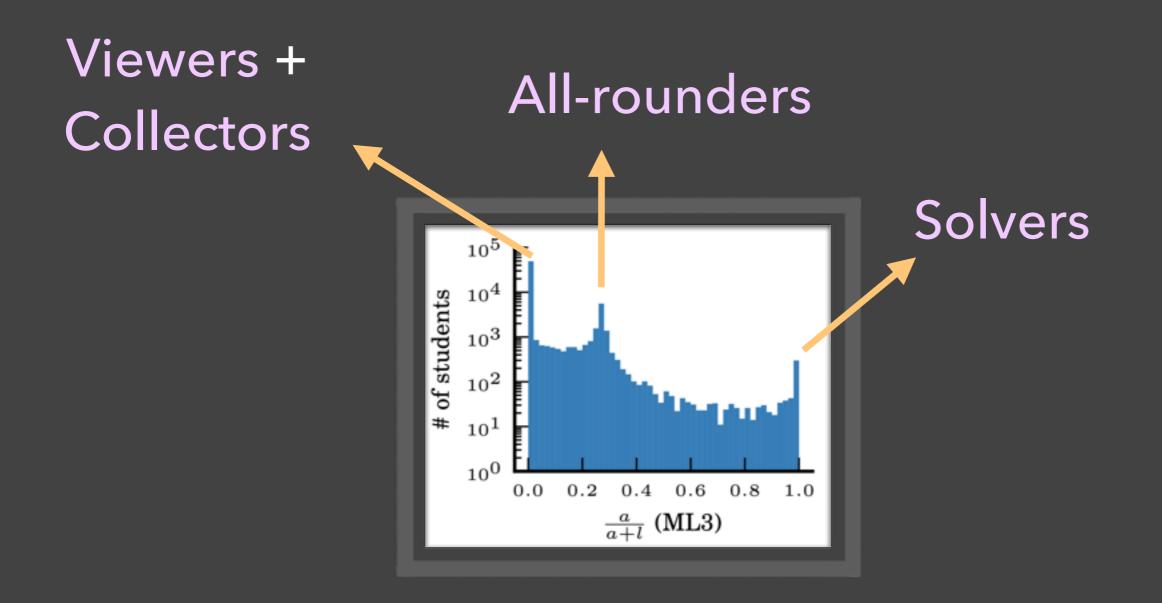
can we classify students by their engagement styles?



Histogram over students' assignment fractions



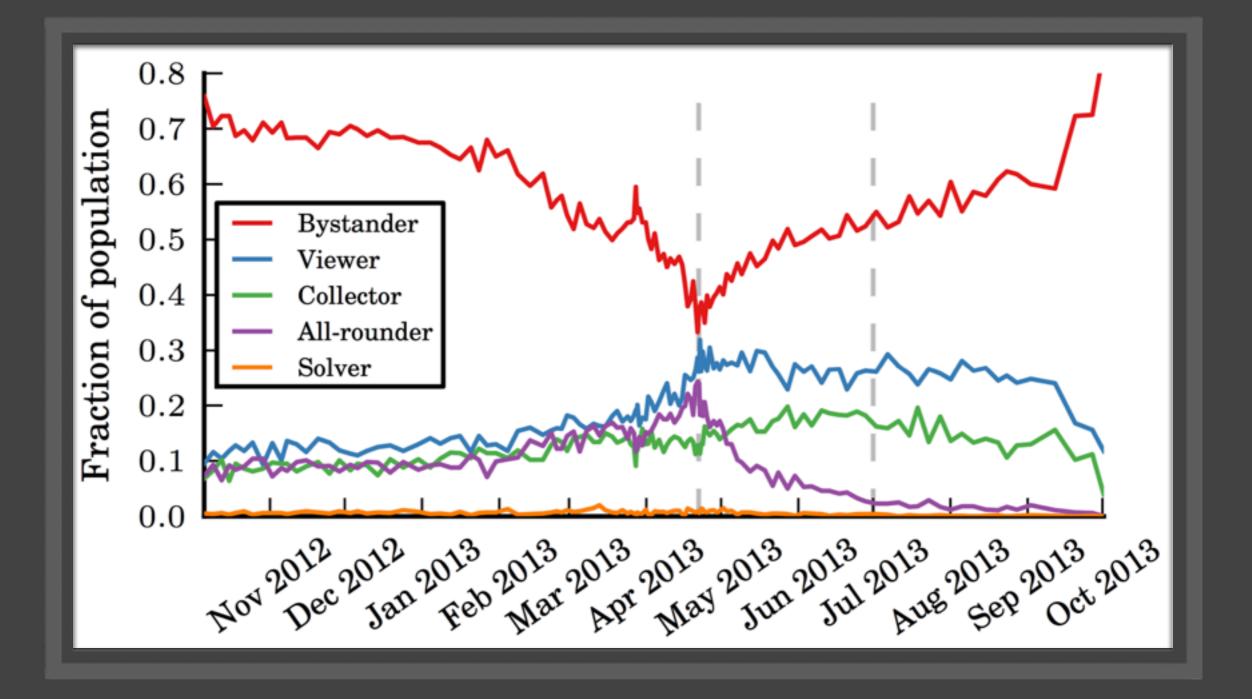
Same for all courses



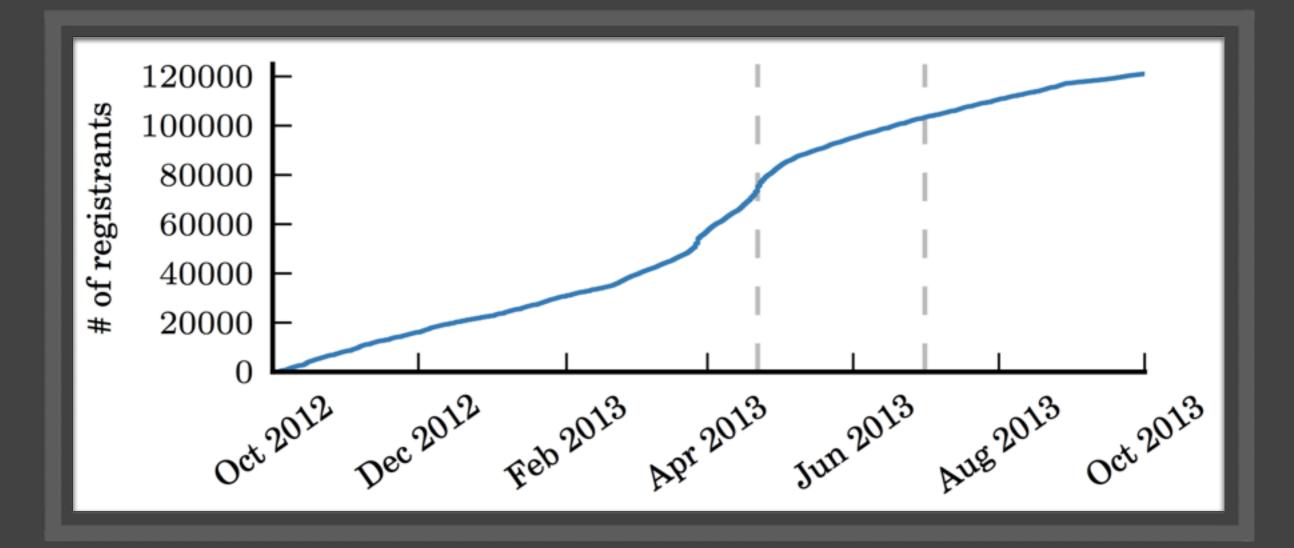
Not many actions: Bystanders

Class	Bystander	Viewer	Collector	All-rounder	Solver
ML1	28,623 (.47)	15246 (.25)	8,850 (.15)	8,067 (.13)	378 (.01)
ML2	27,948 (.49)	13,920 (.21)	7,314 (.11)	9,298 (.19)	550 (.01)
ML3	62,020 (.54)	24,411 (.21)	15,282 (.13)	13,417 (.12)	786 (.01)
PGM1	13,486 (.47)	6,742 (.23)	6,147 (.21)	2,365 (.08)	25 (.00)
PGM2	22,767 (.62)	6,689 (.18)	5,727 (.16)	1,507 (.04)	116 (.00)
PGM3	15,920 (.61)	4,816 (.19)	3,772 (.15)	1,287 (.05)	157 (.01)

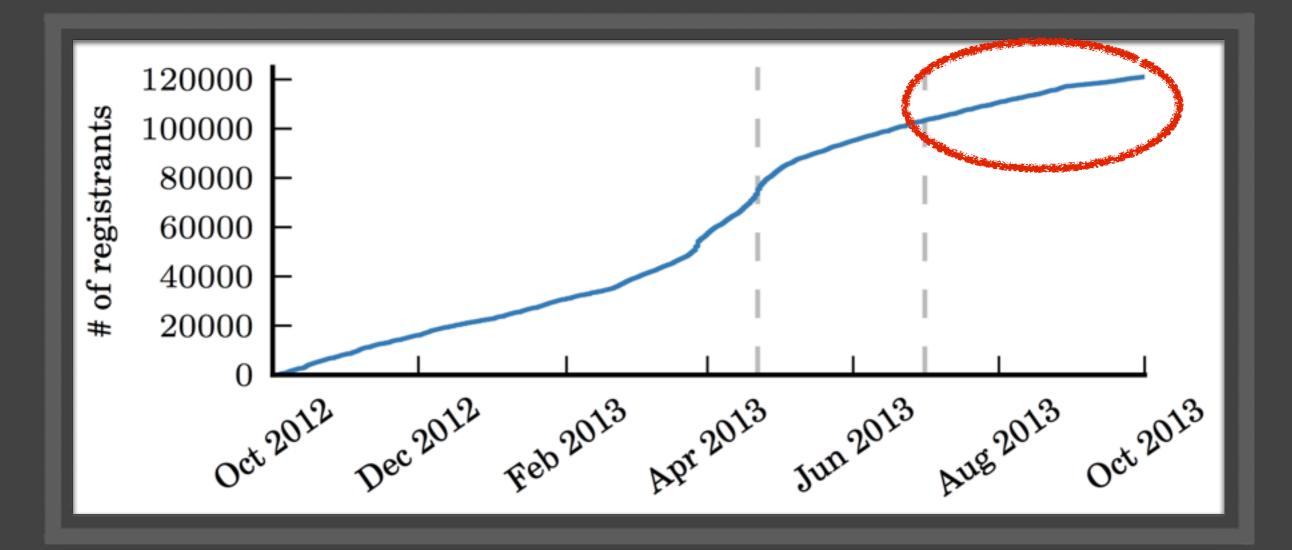
Number (fraction) of students of each style



eventual engagement style vs. registration time



we find a large fraction of *archaelogists*, students who register after the class ends

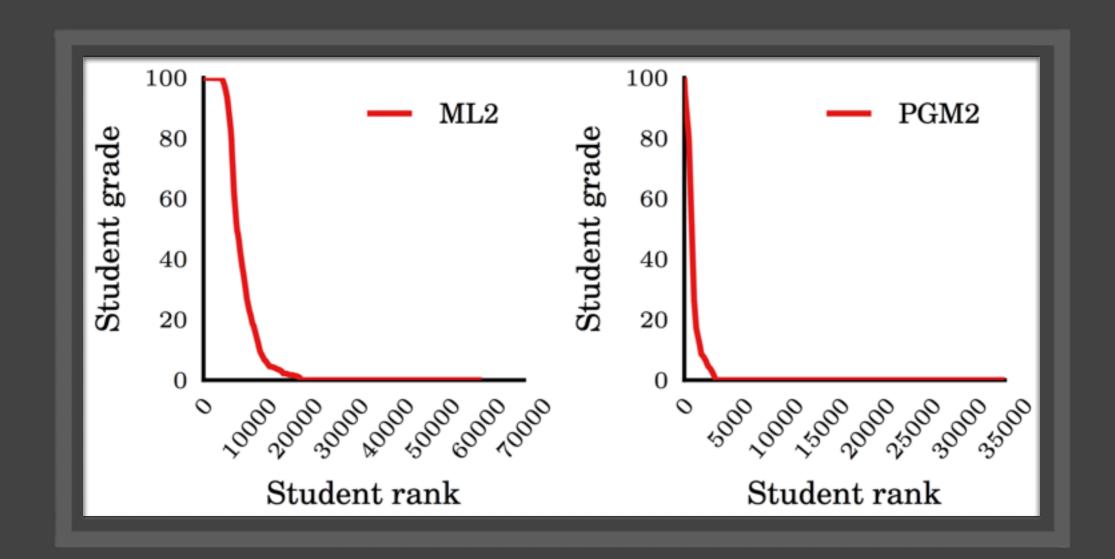


we find a large fraction of *archaelogists*, students who register after the class ends

engaging with massive online courses

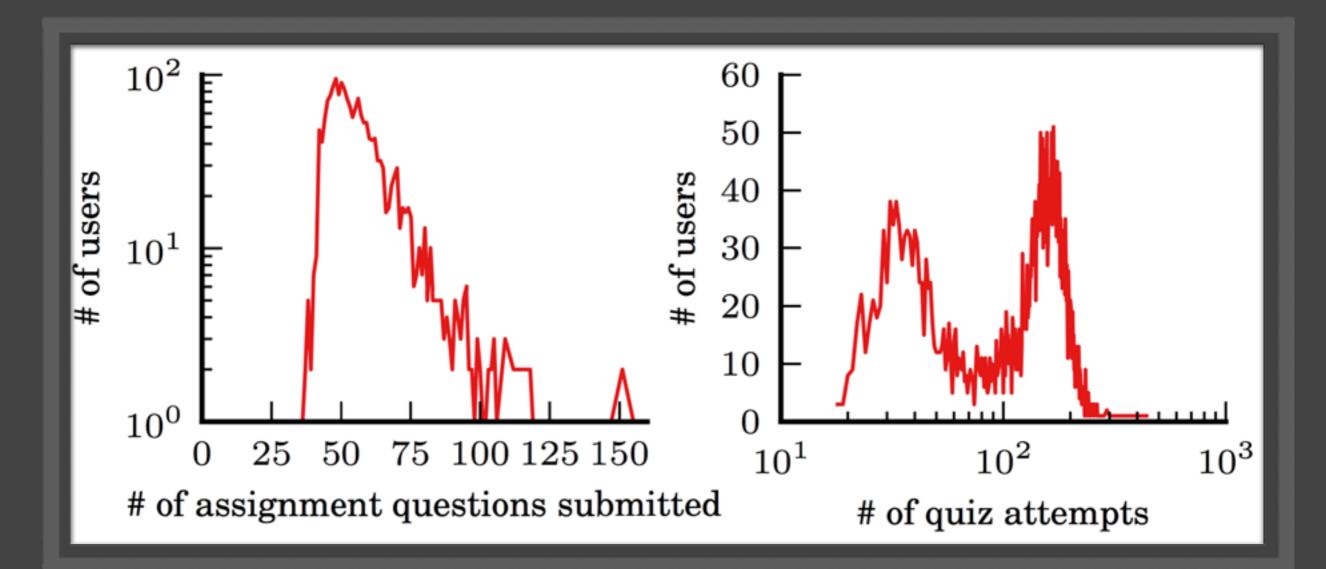
- 1. participation
- 2. performance
- 3. interaction
- 4. intervention

performance



Grades received

performance



Distribution of activities for high-achievers

engaging with massive online courses

- 1. participation
- 2. performance
- 3. interaction
- 4. intervention

General Discussion	
Subscribe for email updates.	0
General discussion about the course. Please read our forum posting policies before posting or starting a new thread.	Search
Sub-forum La	atest Activity
Introductions Make your introductions to your other course mates Co	nnect on Linkedin (14 hours ago)
All Threads Start new thread Top thread	ds Last updated Last created
Is it "theta-transpose * x" or "X * theta"? STAFF REPLIED · Started by Tom Mosher COMMUNITY TA · Last post by Lewis Belcher (3 days ago)	26 19 1547 points posts views
Making your equations look nice in forums made easy. Some Latex examples starff REPLIED · Started by Patrick Campbell · Last post by Valentin Fedulov (5 days ago)	11 8 277 points posts views
# HONOR CODE: Please do not post code related to programming assignments STAFF REPLIED · Started by Richard Creamer COMMUNITY TA · Last post by Eric Borts (5 days ago)	0 5 98 points posts views
Found an error? Check the wiki for errata first Started by Tom Mosher COMMUNITY TA - Last post by Tom Mosher COMMUNITY TA (a month ago)	1 1 250 point post views
Data set that won't fit in memory Started by Lavi Avigdor - Last post by Lavi Avigdor (37 minutes ago)	0 1 1 points post view
Some questions about manifold learning algorithm Started by 杜俊楠 · Last post by 杜俊楠 (3 hours ago)	0 1 7 points post views
What will be your practical application with machine learning? STAFF REPLIED · Started by Rob van Putten · Last post by Helio Perroni Filho (8 hours ago)	11 27 445 points posts views

the discussion forums provide a mechanism for students to interact with each other

what types of students are active on the forums?

how are the forums being used?

	Bystander	Viewer	Collector	All-rounder	Solver
P(S F)	0.106	0.277	0.192	0.408	0.017

P(S|F): given forum usage, which engagement style

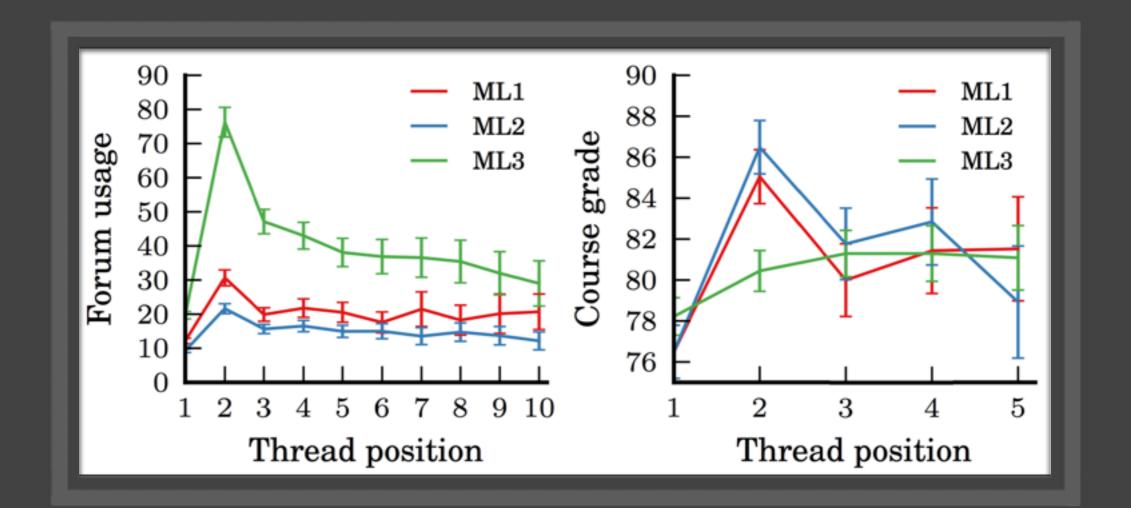
all-rounders and viewers make up most of the forums

	Bystander	Viewer	Collector	All-rounder	Solver
P(S F)	0.106	0.277	0.192	0.408	0.017
P(F S)	0.050	0.334	0.369	0.894	0.648

P(S|F): given forum usage, which engagement style P(F|S): given style, likelihood of forum usage

all-rounders and viewers make up most of the forums

90% of all-rounders are on the forums!



less-active, lower-graded students start threads, more-active, higher-graded students respond

consistent with q&a usage

engaging with massive online courses

- 1. participation
- 2. performance
- 3. interaction
- 4. intervention

can we increase engagement?

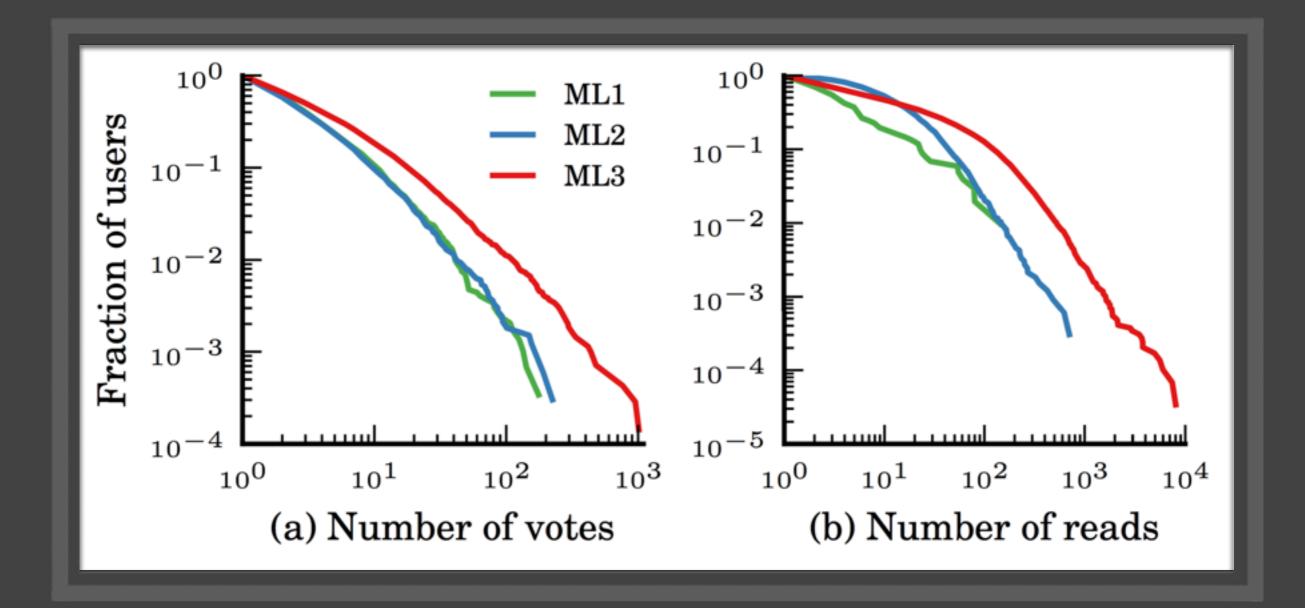
we designed and implemented a badge system to increase forum engagement on ML3

(Thanks to Pamela Fox and Norian Caporale-Berkowitz for the implementation help!) 28

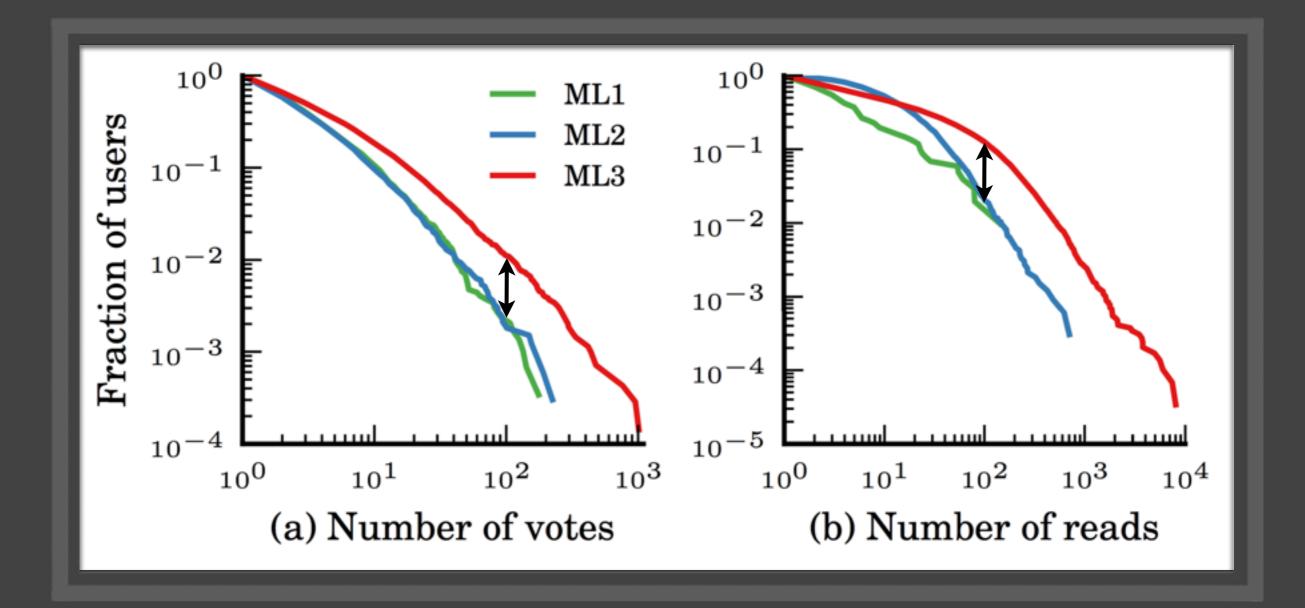
Badge name	Action	Category	Criteria	Num Badges
Supporter	Votes	Cumulative	Awarded once a user votes 3/15/40/100 times	4 (BSGD)
Reader	Reading threads	Cumulative	Read 10/30/70/200 threads	4 (BSGD)
Good/Great/ Awesome/ Incredible Reply	Quality replies	Great Achievement	Awarded for contributing a high-quality reply (5/10/25/100 upvotes)	4 (BSGD)
Good/Great/ Awesome/ Incredible Thread	Quality threads	Great Achievement	Awarded for contributing a high-quality thread (5/10/25/100 upvotes)	4 (BSGD)
Contributor	"Good" replies	Cumulative Great Achievement	Contributing 3/6/10/25 good reply (where good = 3 upvotes)	4 (BSGD)
Conversation Starter	"Good" threads	Cumulative Great Achievement	Contributing 3/6/10/25 good threads (where good = 3 upvotes)	4 (BSGD)
Community Member	Join class	First-time	Awarded when user joins class (as intro to badges)	1
Forum Newbie	Any	First-time	Awarded once a user takes any action in the forums	1
Early Bird	Vote/Post/Thread	Activity	Active on forums in first two weeks	1
All-Star	Vote/Post/Thread	Activity	Being active in all weeks	1

did the badges have an effect?

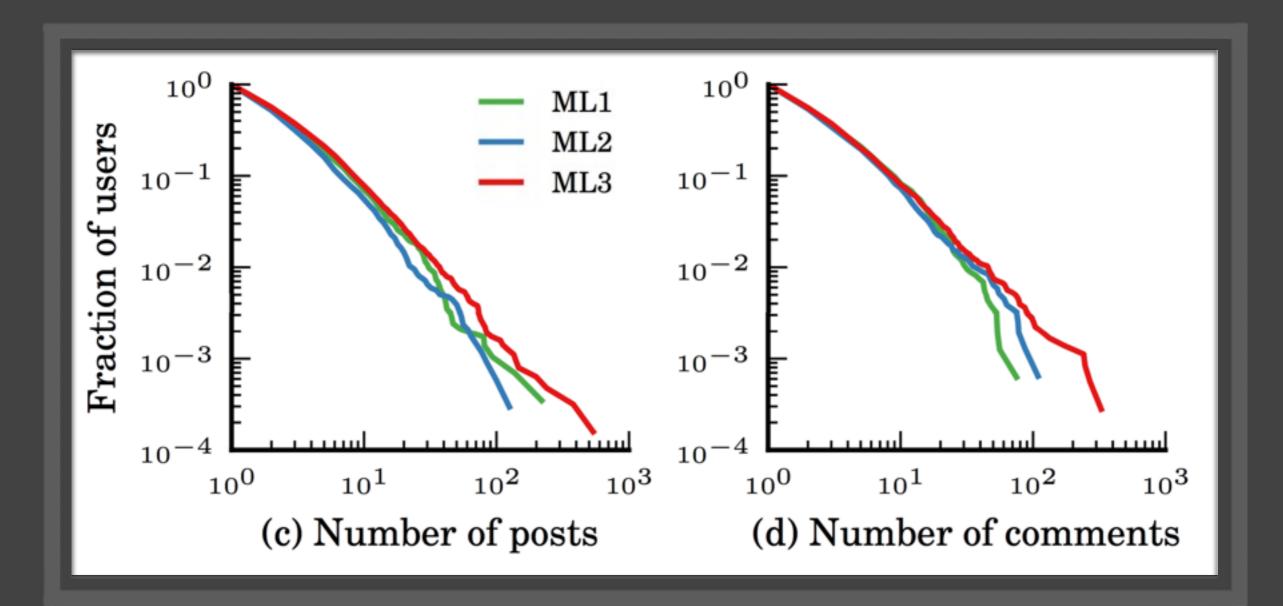
implemented badges on ML3, compare observationally with previous runs ML1 and ML2



5x more likely to get to 100 votes/reads!



5x more likely to get to 100 votes/reads!



no qualitative difference in posts/comments no badges on these actions!

badgified dimensions \Rightarrow 5-fold increase in engagement unbadgified dimensions \Rightarrow no qualitative difference

> not a true experiment, but very strong observational evidence of badge effect

engagement can be increased in targeted ways!

a true experiment: variation in badge presentation what gives badges their power?

> compare different badge presentations, measure which have strongest effects

three experiments:

top bylines
 thread bylines
 badge ladder

factorial design (users randomly assigned to 1 of 8 (=2³) buckets)

experiment 1: top badge byline

oursera		Courses Ash Ton 🕶	heade
Stanford 🕷	achine Learning Andrew Ng		
Image: Course Schedule	Announcements Week 3 Updates, and a Reminder about Due Dates Helo al! I hope everyone has been enjoying the course and learning a lot The video lectures, programming exercises, and review questions for Week 4 are now available. The due date for the second programming exercise is quickly approaching. It is due this Sunday, April 6, 2014 at 11:59 PM Pacific Time. Remember that you can use our https://class.coursera.org/mI-005/class. Antrew Ng	Upcoming Deadlines Review Questions VL Logistic Regression But 6 Apr 2014 11:59 PM PDT VL Regularization But 6 Apr 2014 11:59 PM PDT Programming Assignments Logistic Regression But 6 Apr 2014 11:59 PM PDT Use IRC? Join #mi-class on	
Course Information		irc.freenode.net to chat with your classmates.	
Tutoring	Week 2 Updates, and a Reminder about Due Dates		
Course FAQ	Hello all	Recent Discussions	
Octave // Matlab Tutorial	I hope everyone has been enjoying the course and learning a lot! New material has just been posted, which you can see at https://class.coursera.org/ml-005/lecture/index.	Problem with submitting Part 2 oneVsAll for Programming Exercise 3 Last port by Coin Beckingham	
Course Wiki 🛤	The due date for the first programming exercise is quickly approaching. It is due this Sunday,	COMMUNITY TA: (Just now)	
Join a Meetup≌	March 30, 2014 at 11:59 PM Pacific Time. All submissions after this time will be penalized by 20%. If you miss the deadline, we encourage you to try to submit anyway, because you can	costFunctionReg Problem Last post by Maria Polugodina (a minute ago)	
Help Articles Course Materials Errors Technical Issues	still get some credit for it. In addition, our course wiki is now open! The course wiki is a place where you can share useful advice with your classmates about the course content, review quizzes, and programming exercises. Andrew Ng	gradientDescent.m not submitting Last post by AKSHWY AGARWAL (3 minutes app) Regularization results seem correct but submit fails Last post by Sha Haim (3 minutes app)	
	Mon 24 Mar 2014 10:00 AM PDT	unable to see past programming exercises Last post by Coln Beckingham	

experiment 1: top badge byline

Control:

Courses Ash Ton 🔻

Treatment:



experiment 2: thread badge bylines

↑ 10 ↓ ·	fan .	
Ψ 10 Ψ .	range	
Rich	and Creamer Commonweak - 24 days ago %	
Thank you, ✿ 0 � · fl	Patrick! I'll take a look as I've been looking for a quick tutorial on LaTex. ag	
		+ Comment
2 Rob	ert Tullo - 24 days ago %	
very cool. ink is here	I did find this a few days ago for when I needed to post.	
↑1 ↓·1	ag	_
	Patrick Campbell - 23 days ago %	
	ice linkt ► 0 I I - flag	
		+ Comment
Jim Melt	on community and - 11 days ago %	
	eful tool for creating LaTex equations is one of the on-line LaTex equation editors. I like this one: .codecogs.com/latex/eqneditor.php	
You can us \$\$' delimit	e a GUI to get you started with the formula, and when you have it the way you want it, just copy and paste it ers.	t between
-	re using Rich Text mode, you can click the "Math" button and paste your equation there. One caveat: Your line (just delete any new-line characters so it all runs together)	LaTex must
↑ 0↓·1	ag	_
		+ Comment
	ntin Fedulov - 5 days ago %	
2 Vale		

Posts are annotated with author name and timestamp

experiment 2: thread badge bylines

Result	
$a \ (1, 2,)$	
↑ 10 ↓ - flag	
and the second design of the	
Richard Creamer cosmuners - 24 days ago %	
Thank you, Patrick! I'll take a look as i've been looking for a guick tutorial on LaTex.	
♠ 0 ↓ · fag	
	+ Comment
Robert Tulio - 24 days ago %	
very cool. I did find this a few days ago for when I needed to post.	
link is here	
↑ 1 ↓ · fag	
Patrick Campbell - 23 days ago %	
Nice Ink	
↑ 0 ↓ · flag	
	+ Comment
Jim Melton cowwarmink - 11 days ago %	
Another useful tool for creasing LaTex equations is one of the on-line LaTex equation editors. I like this one:	
http://www.codecogs.com/latex/eqneditor.php	
You can use a GUI to get you started with the formula, and when you have it the way you want it, just copy and	paste it between
'\$\$' delimiters.	
Or, if you are using Rich Text mode, you can click the "Math" button and paste your equation there. One caveat be a single line (just delete any new-line characters so it all runs together)	: Your LaTex must
↑ 0 ↓ · flag	
	+ Comment
	- communit
🖳 Valentin Fedulov - 5 days ago %	
$\frac{\partial}{\partial \theta}J(\theta) = \frac{1}{m}\sum_{i=1}^{m} (\theta^{T}x^{i\theta} - y^{i\theta})x_{j}^{i\theta}$	
00j (c) m Autol (c) m (c) / / //	
Just checking - it works!	
Just checking - it works! Don't forget add empty line between formula and other text. And a cheatsheet:	

Posts are annotated with author name and timestamp

experiment 2: thread badge bylines

Control:

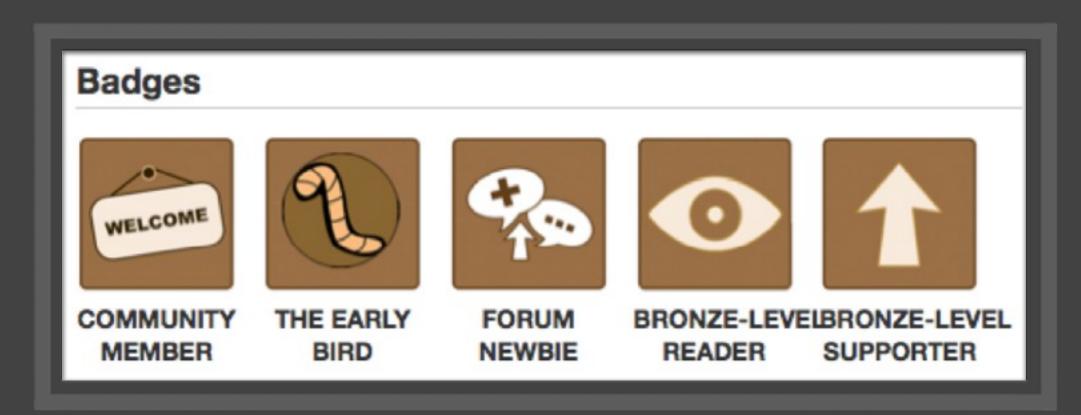
Patrick Campbell · 23 days ago %

Treatment:

Connorelly 2 1 1 1 · 2 months ago %

experiment 3: badge ladder

Control:



experiment 3: badge ladder Treatment:

Badge Series (2 earned)	_	_	_	_
	BRONZE	SILVER	GOLD	DIAMOND
The Reader To earn the next badge (Silver), you must read 30 threads from your classmates.	•			
The Supporter To earn the next badge (Silver), you must vote on 15 posts that you find interesting or useful.				
The Contributor To earn the next badge (Bronze), you must post 3 replies that your classmates find interesting.			-	
The Conversation Starter To earn the next badge (Bronze), you must start 3 threads that your classmates find interesting.				
Top Posts To earn the next badge (Bronze), you must write a post that gets 5 upvotes from your classmates.			-	

Top byline	Thread byline	Badge ladder
0.095	0.095	0.036

Mann-Whitney rank-sum p-values

Badge ladder most significant Explicit goal-setting helped more than increased social visibility of badges

conclusion

conceptual framework for the quantitative analysis of engagement in moocs

classified users into a taxonomy of engagement styles

designed and implemented a badge system and a randomized experiment that increased forum engagement

thank you!