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

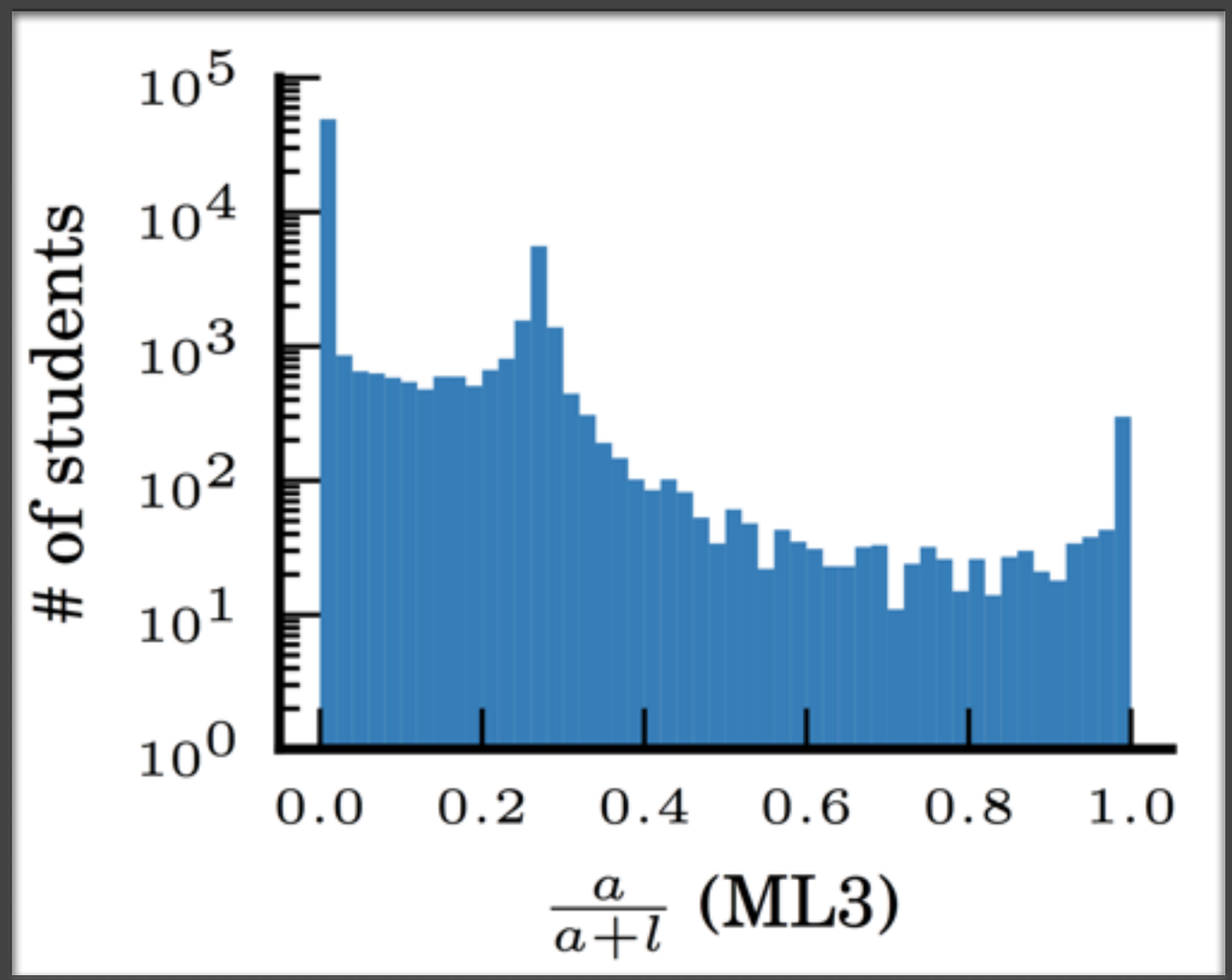
1. participation
2. performance
3. interaction
4. intervention

engagement styles

are different students using moocs differently?

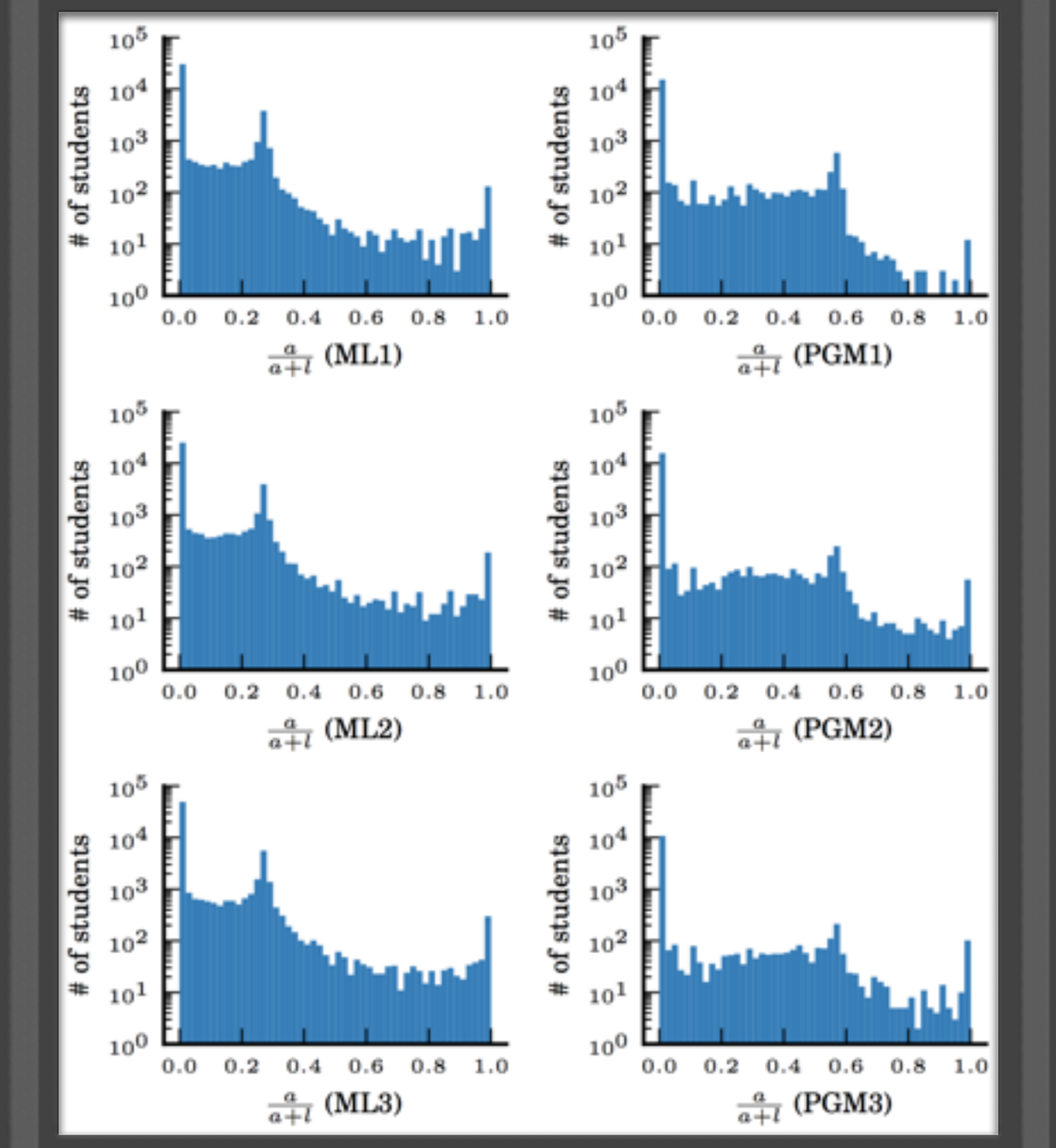
can we classify students by
their *engagement styles*?

engagement styles



Histogram over students' *assignment fractions*

engagement styles



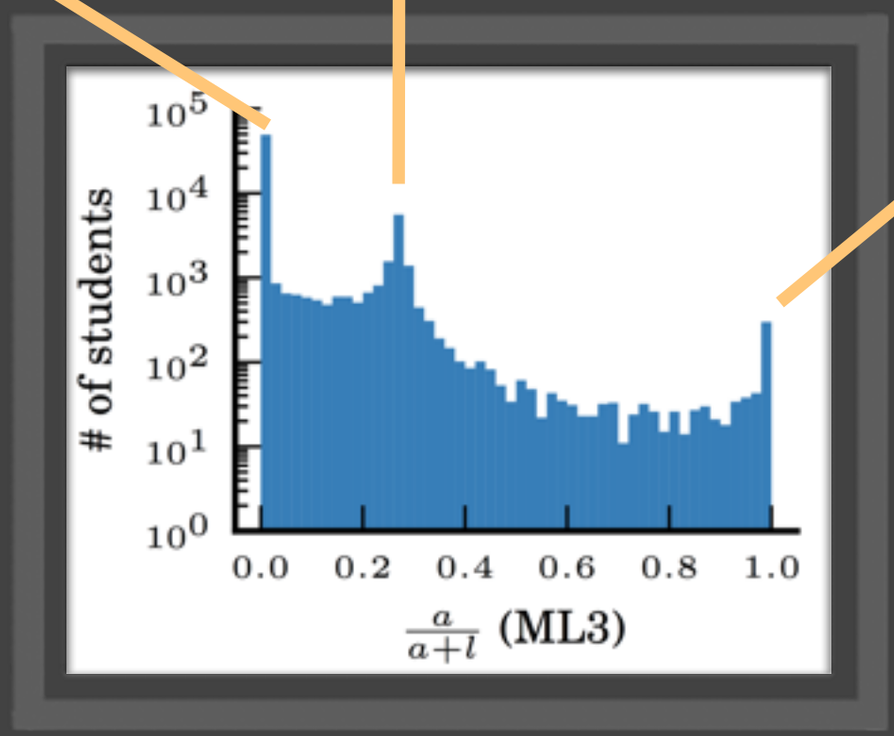
Same for all courses

engagement styles

Viewers +
Collectors

All-rounders

Solvers



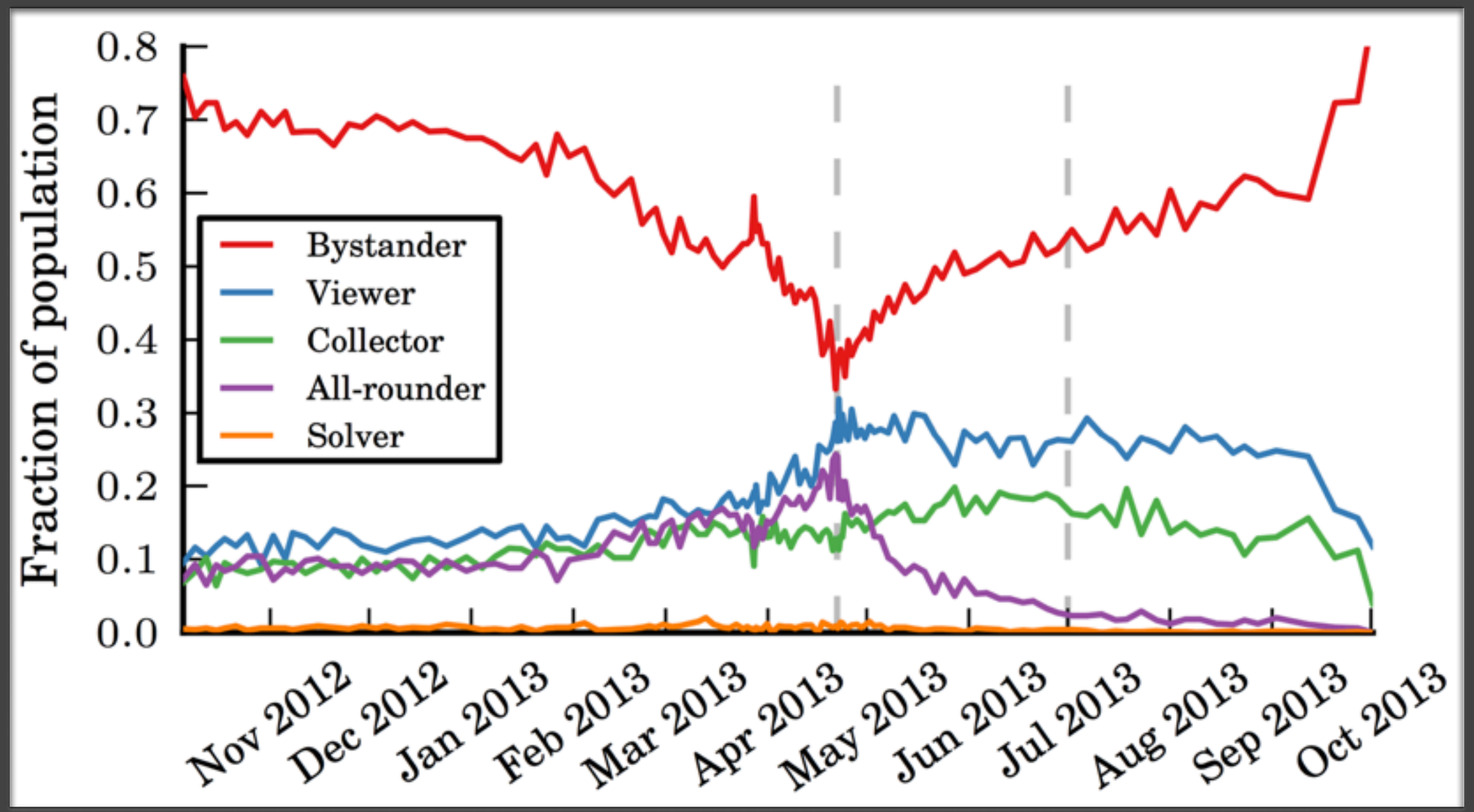
Not many actions: Bystanders

engagement styles

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)

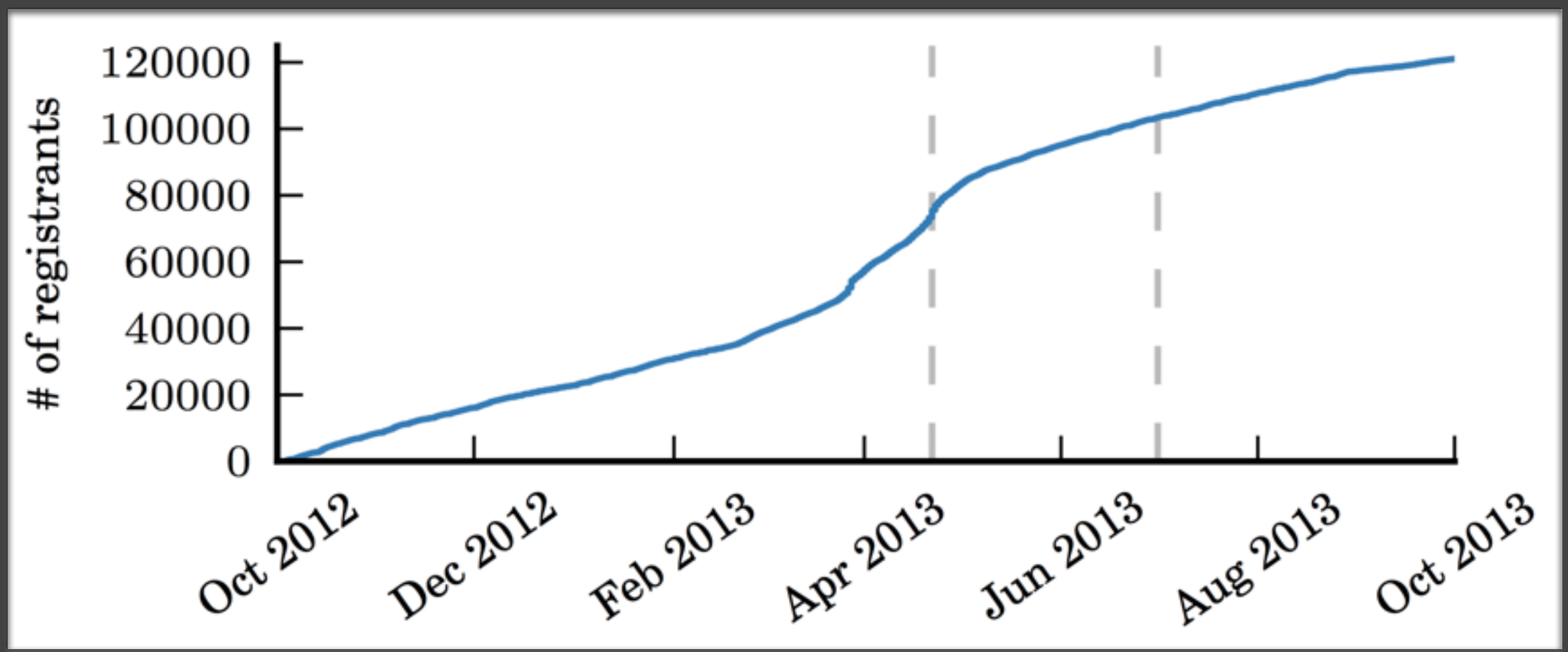
Number (fraction) of students of each style

engagement styles



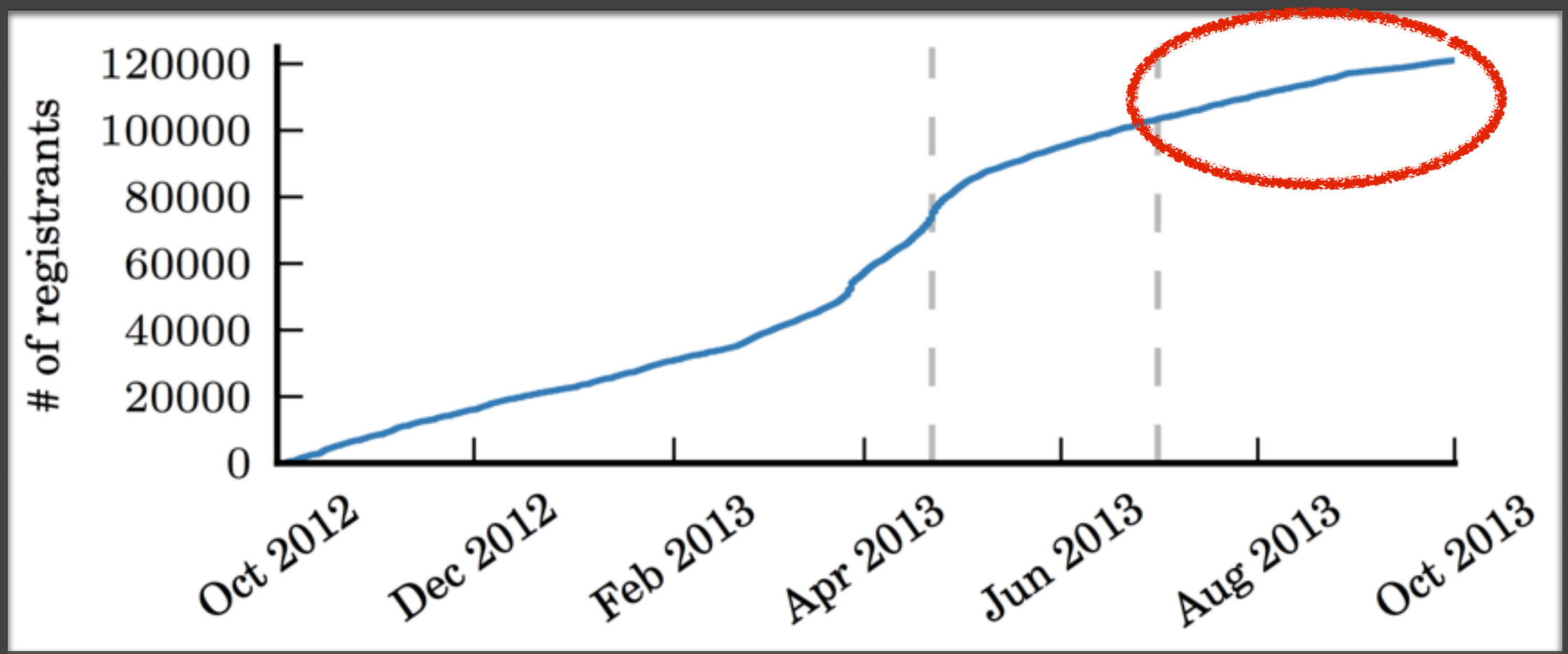
eventual engagement style vs. registration time

engagement styles



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

engagement styles

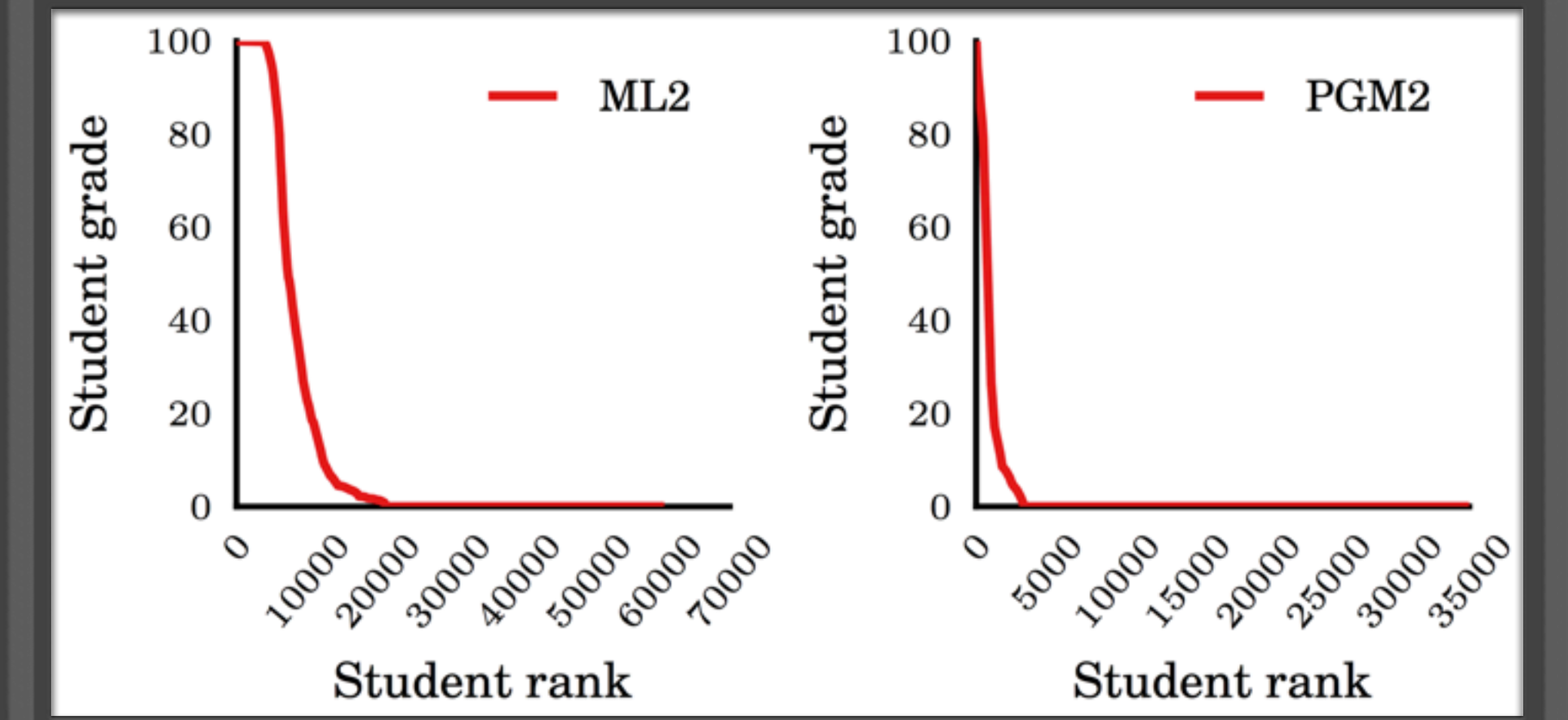


we find a large fraction of *archaeologists*, students who register after the class ends

engaging with massive online courses

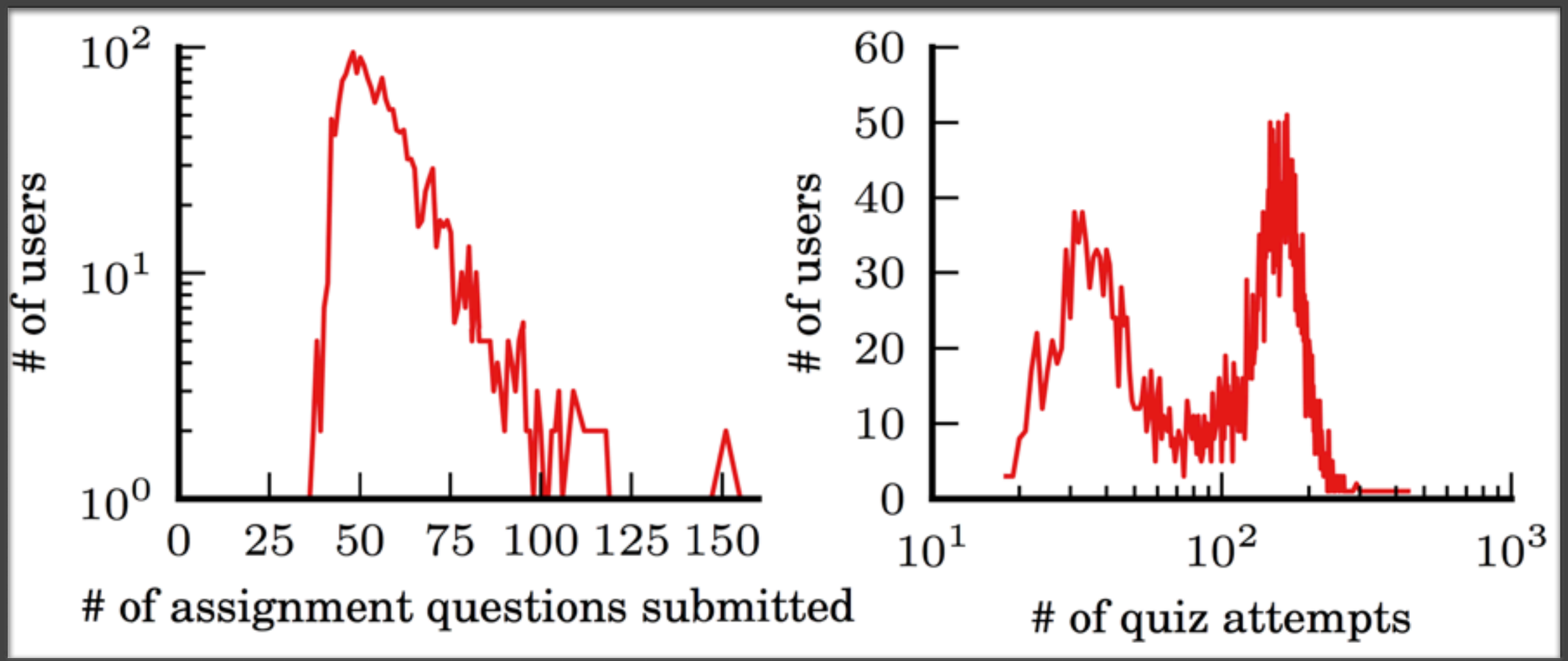
1. participation
2. performance
3. interaction
4. intervention

performance



Grades received

performance

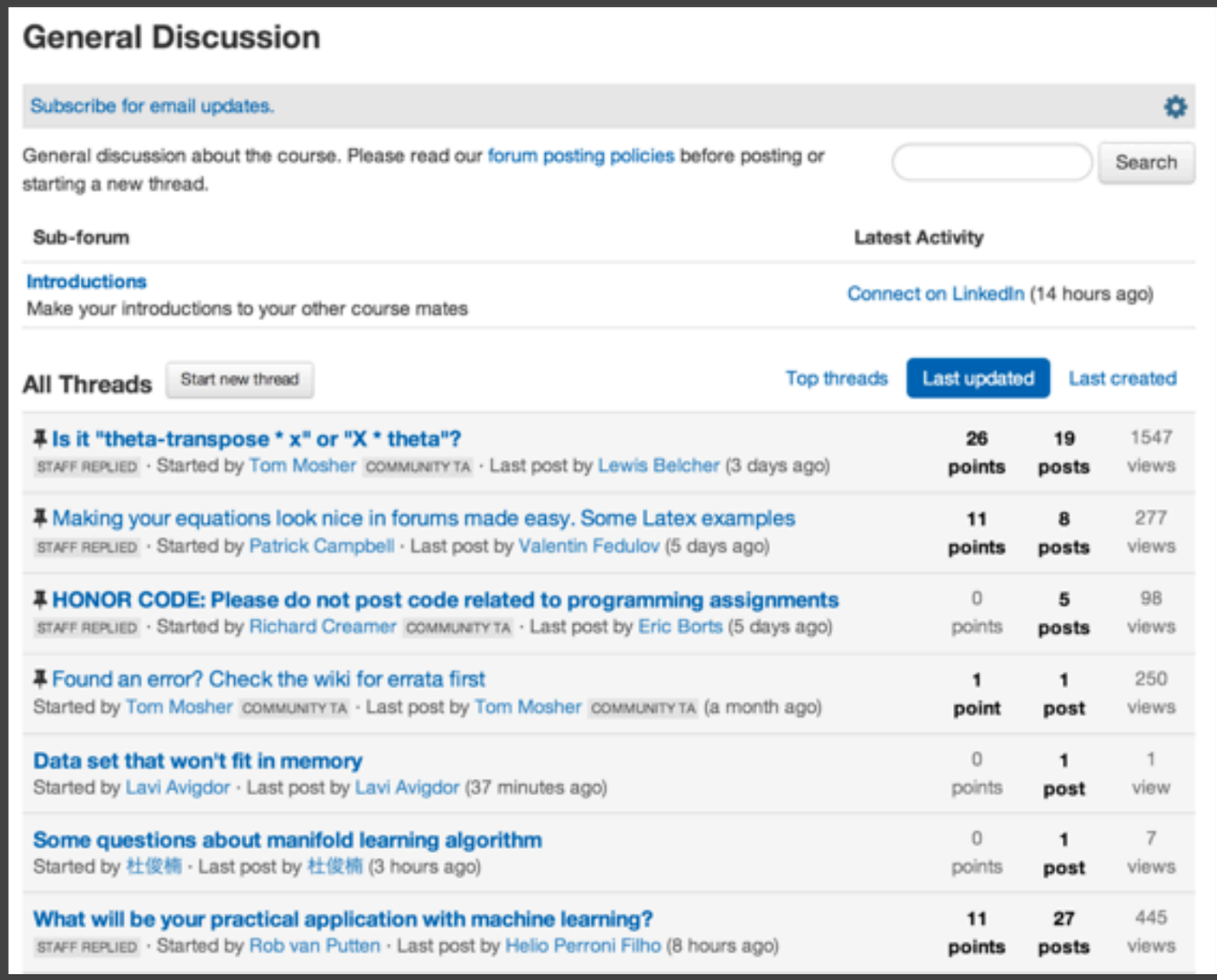


Distribution of activities for high-achievers

engaging with massive online courses

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

interaction



General Discussion

Subscribe for email updates.

General discussion about the course. Please read our forum posting policies before posting or starting a new thread.

Search

Sub-forum	Latest Activity
Introductions Make your introductions to your other course mates	Connect on LinkedIn (14 hours ago)

All Threads Start new thread

Top threads | 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 points**	**19 posts**	1547 views
Making your equations look nice in forums made easy. Some Latex examples STAFF REPLIED · Started by Patrick Campbell · Last post by Valentin Fedulov (5 days ago)	**11 points**	**8 posts**	277 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 points	**5 posts**	98 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 point**	**1 post**	250 views
Data set that won't fit in memory Started by Lavi Avigdor · Last post by Lavi Avigdor (37 minutes ago)	0 points	**1 post**	1 view
Some questions about manifold learning algorithm Started by 杜俊楠 · Last post by 杜俊楠 (3 hours ago)	0 points	**1 post**	7 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 points**	**27 posts**	445 views

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

interaction

what types of students are active on the forums?

how are the forums being used?

interaction

	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

interaction

	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

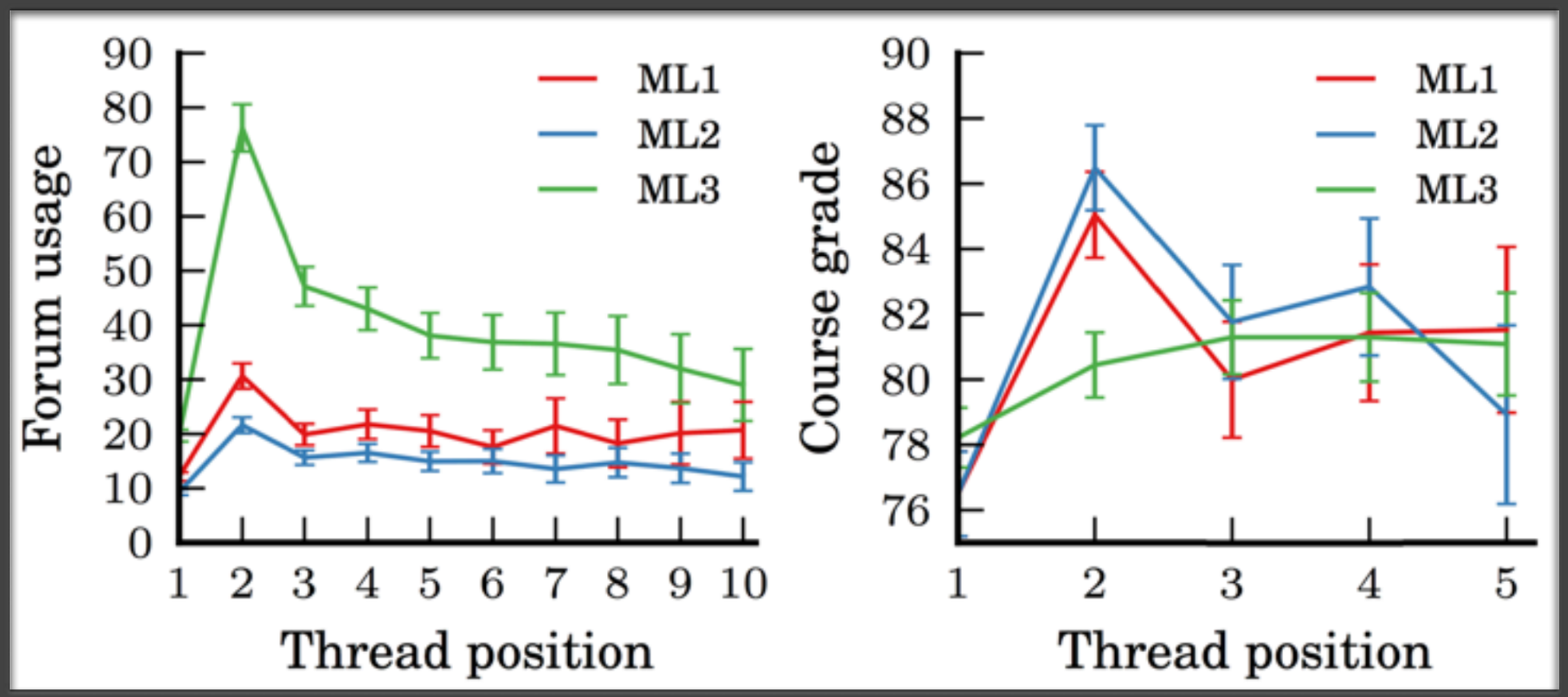
$P(S|F)$: given forum usage, which engagement style

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

all-rounders and viewers make up most of the forums

90% of all-rounders are on the forums!

interaction



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

intervention

can we increase engagement?

intervention

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!)

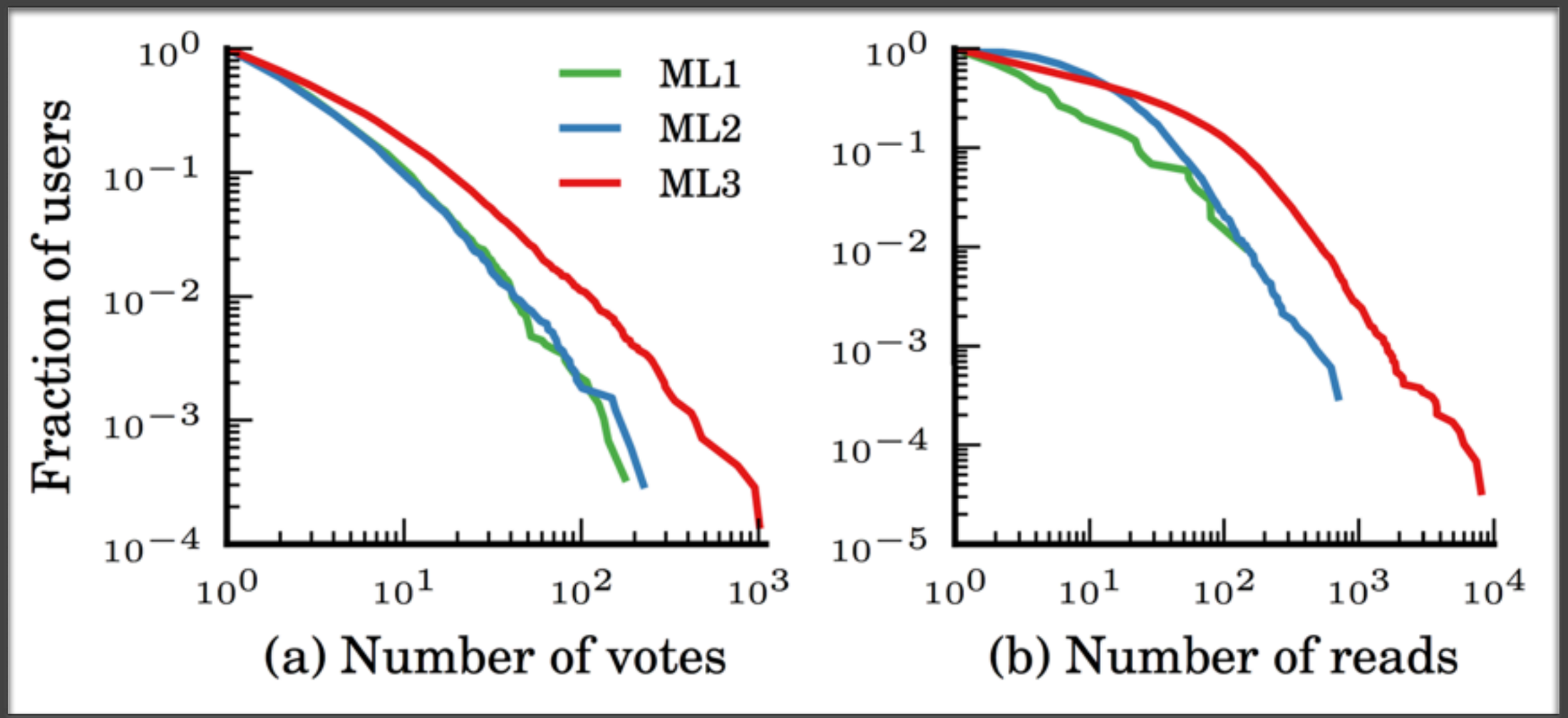
Badge name	*Action*	*Category*	*Criteria*	*Num Badges*
Supporter	Votes	**Cumulative**	Awarded once a user votes 3/15/40/100 times	4 (BSGD)
Reader	Reading threads	**Cumulative**	Read 10/30/70/200 threads	4 (BSGD)
Good/Great/ Awesome/ Incredible Reply	Quality replies	**Great Achievement**	Awarded for contributing 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

intervention

did the badges have an effect?

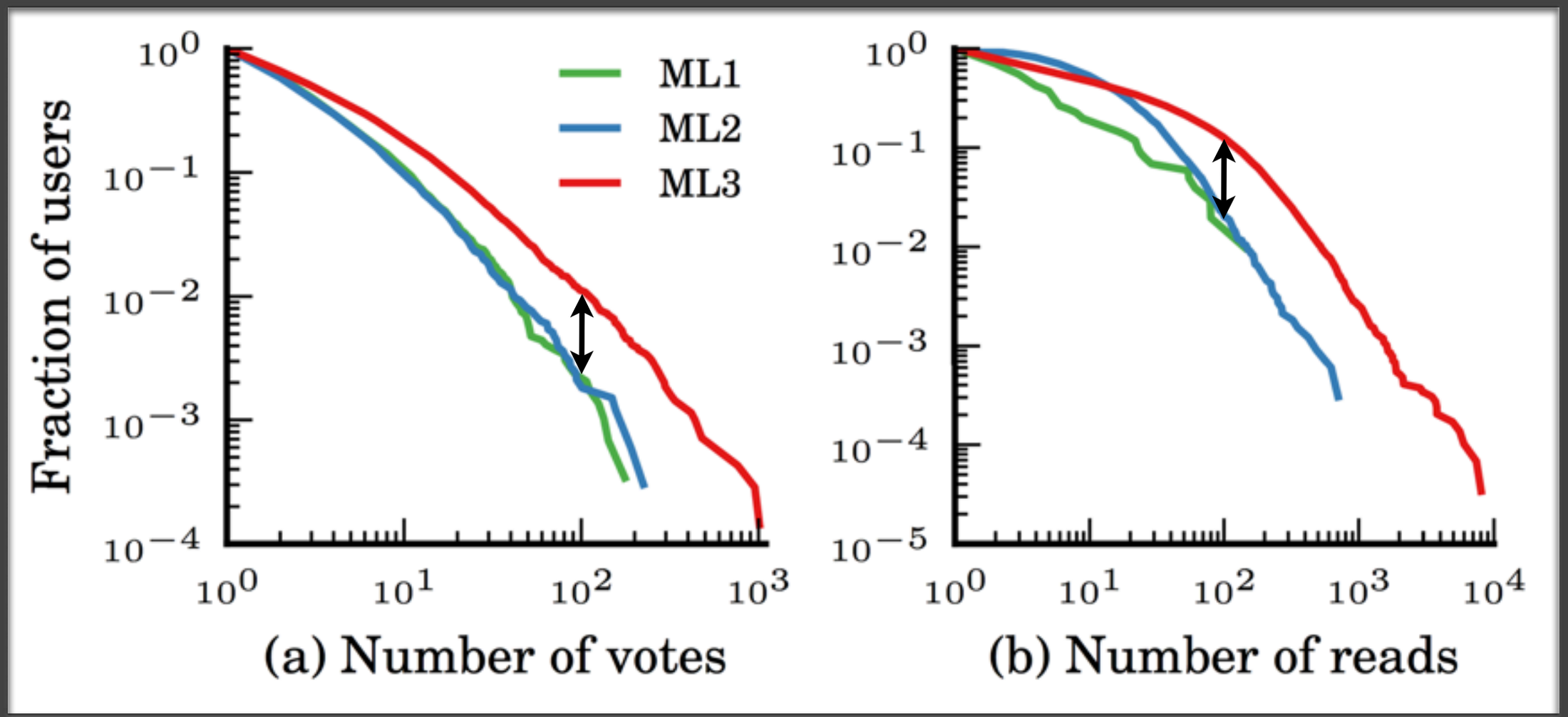
implemented badges on ML3, compare observationally with previous runs ML1 and ML2

intervention



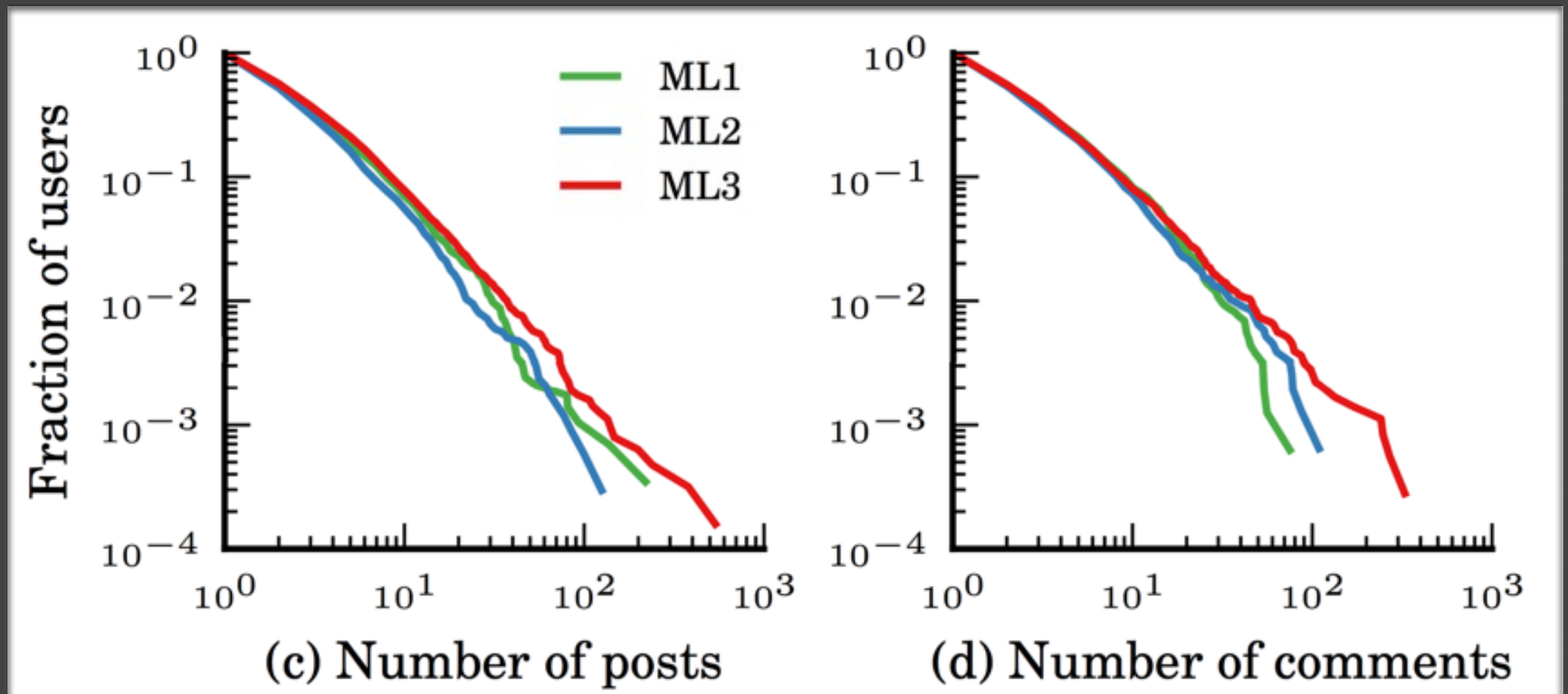
5x more likely to get to 100 votes/reads!

intervention



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

intervention



no qualitative difference in posts/comments

no badges on these actions!

intervention

badgified dimensions ⇒ 5-fold increase in engagement

unbadgified dimensions ⇒ no qualitative difference

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

engagement can be increased in targeted ways!

intervention, part 2

a true experiment: variation in badge presentation

what gives badges their power?

compare different badge presentations,
measure which have strongest effects

intervention, part 2

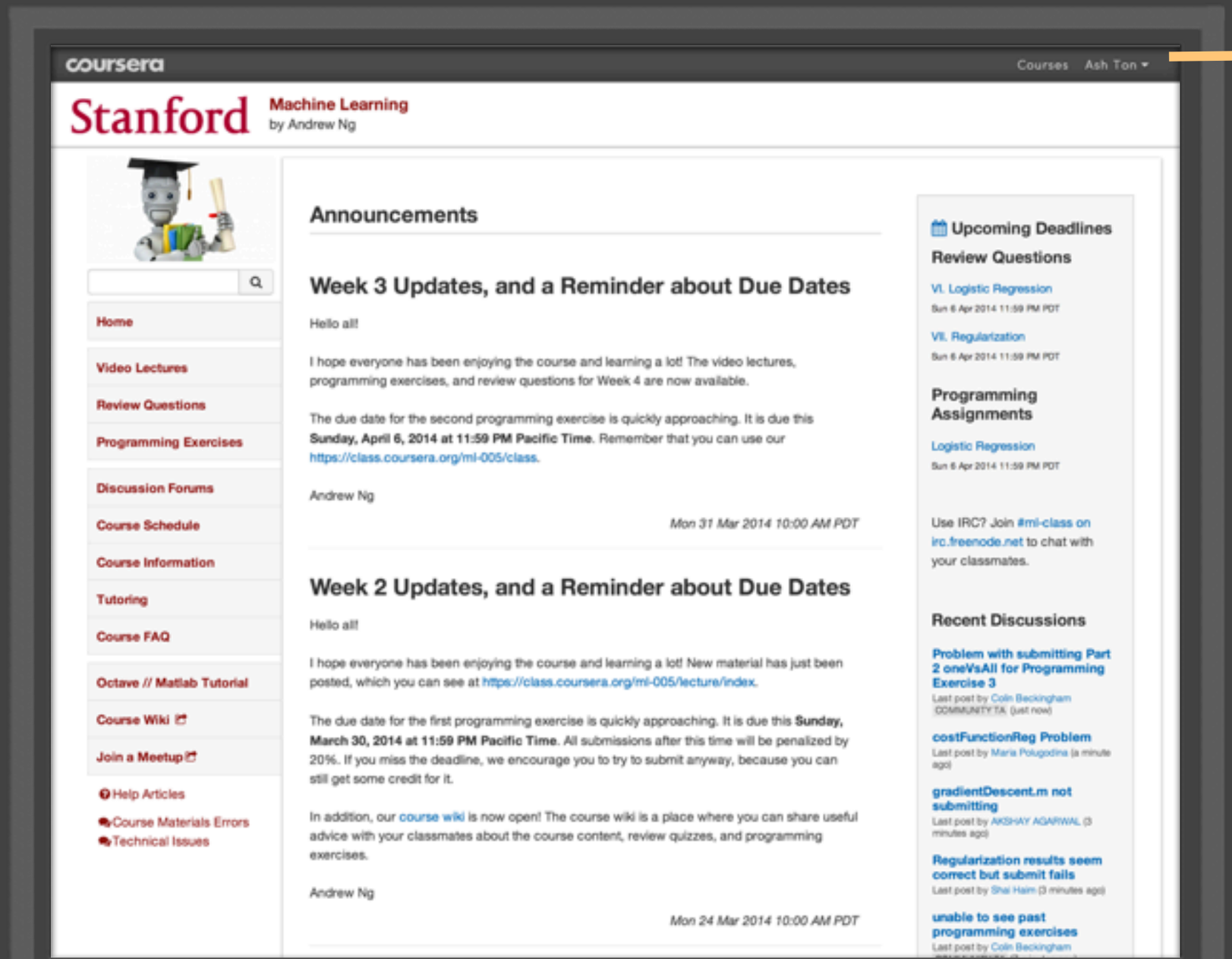
three experiments:

1. top bylines
2. thread bylines
3. badge ladder

factorial design (users randomly assigned to 1 of 8 ($=2^3$) buckets)

intervention, part 2

experiment 1: top badge byline

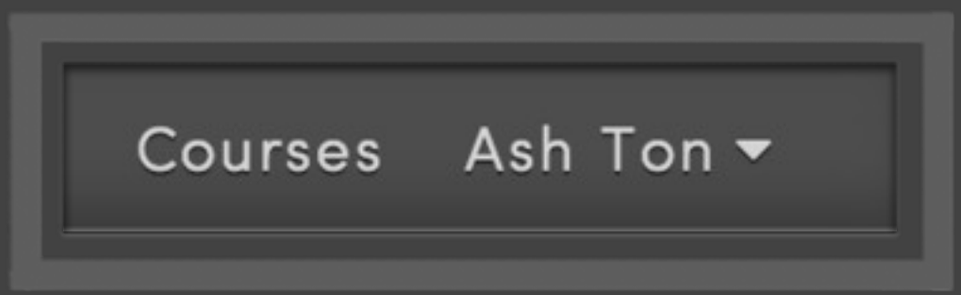


header

intervention, part 2

experiment 1: top badge byline

Control:



Treatment:

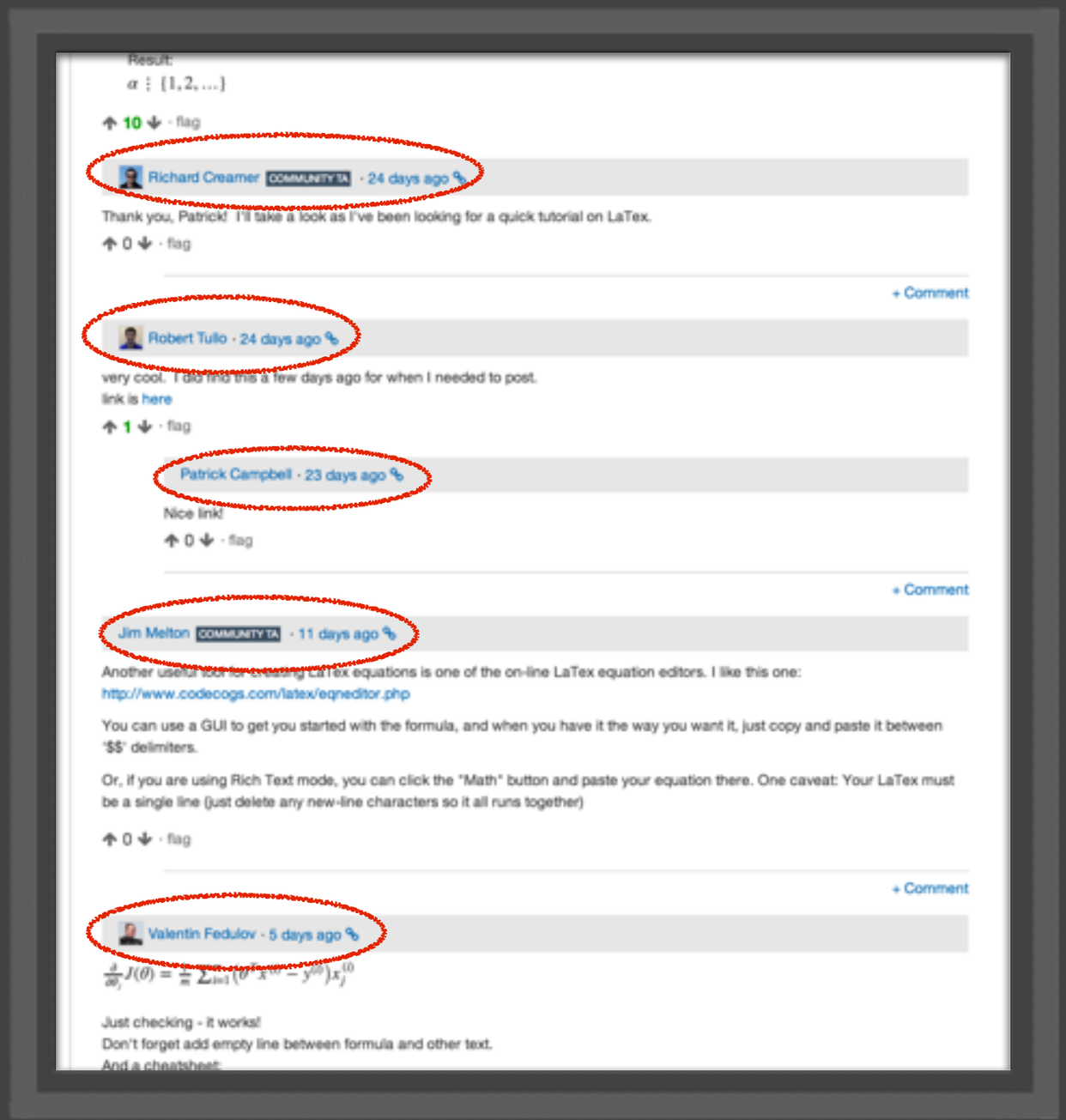
intervention, part 2

experiment 2: thread badge bylines

Posts are annotated with author name and timestamp

intervention, part 2

experiment 2: thread badge bylines

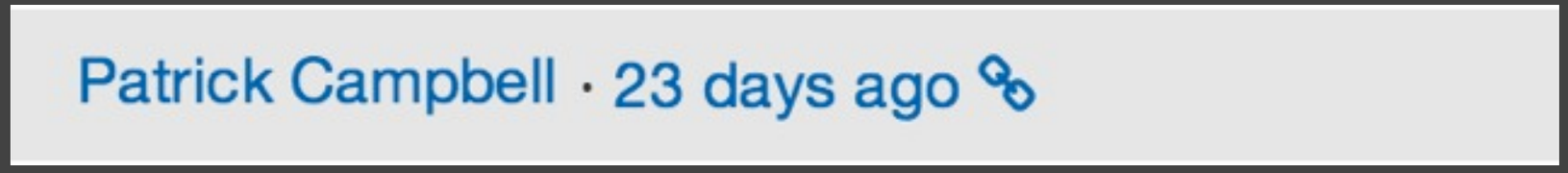


Posts are annotated with author name and timestamp

intervention, part 2

experiment 2: thread badge bylines

Control:

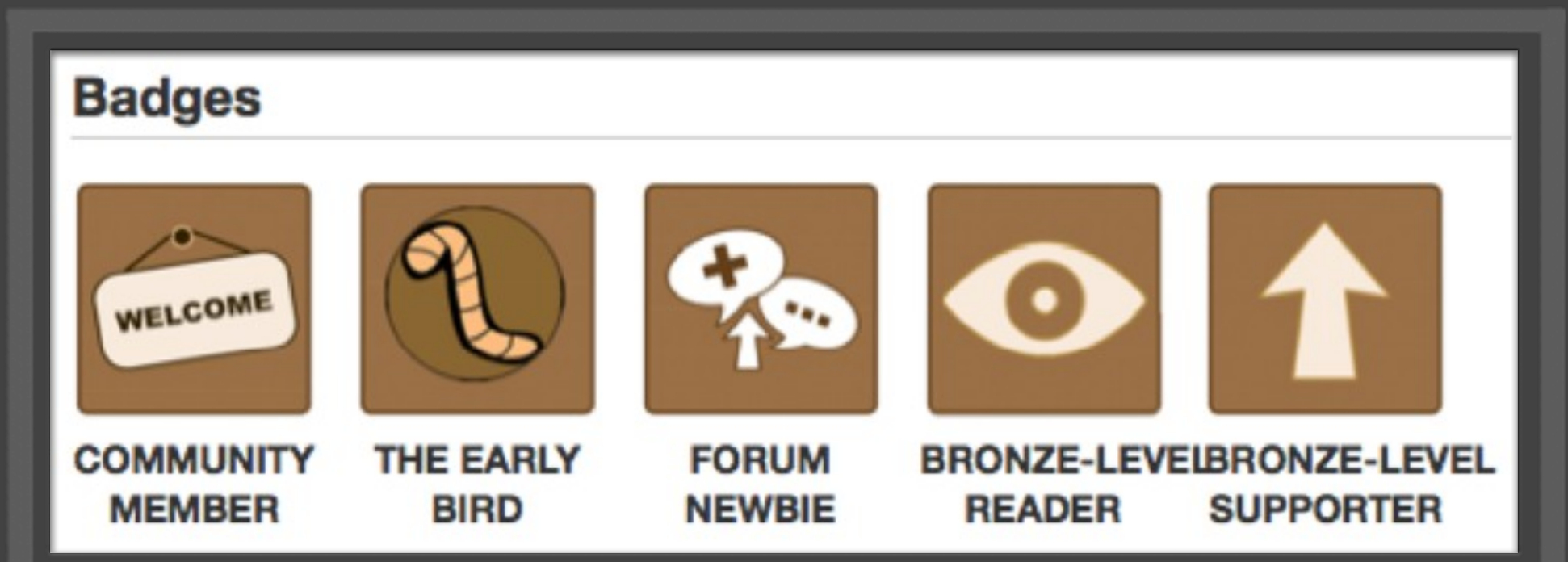


Treatment:

intervention, part 2

experiment 3: badge ladder

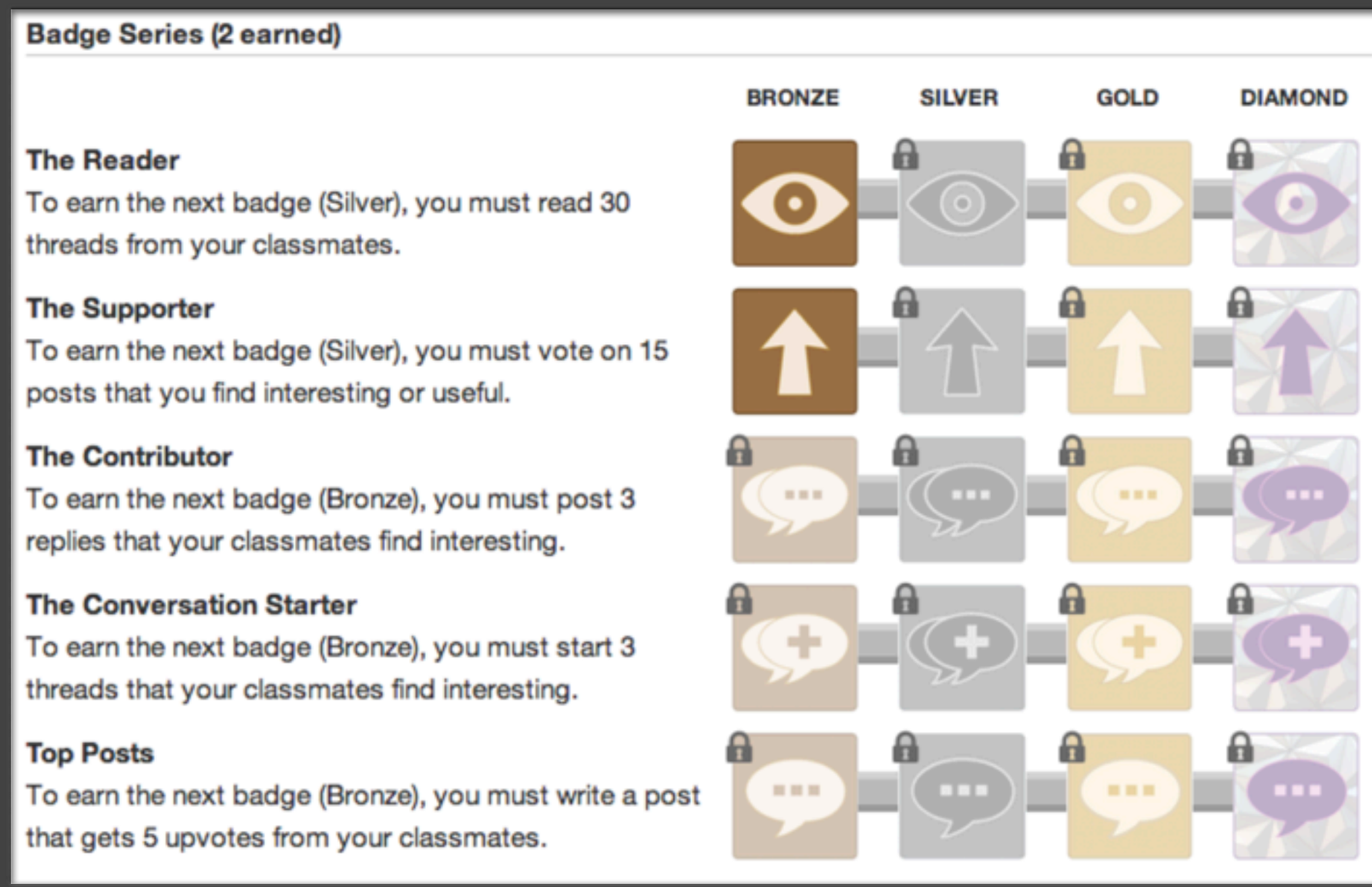
Control:



intervention, part 2

experiment 3: badge ladder

Treatment:



intervention, part 2

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!