Peer Mediated Testing

Igor Labutov, Kelvin Luu, Thorsten Joachims, Hod Lipson
Cornell University
What is Peer Mediated Testing?
Peers generate assessment medium
The Assessment Medium
The Assessment Medium
The Assessment Medium
The Assessment Medium
The Assessment Medium
The Assessment Medium

Classroom: Blue Book, edX, Coursera

Forum: Piazza, Stack Exchange
The Assessment Medium

Classroom

- Structured
- Expert-curated (well calibrated?)
- Expensive

Forum

- edX
- Coursera
- piazza
- Stack Exchange
The Assessment Medium

Classroom
- Structured
- Expert-curated (well calibrated?)
- Expensive

Forum
- Free-form
- Student-driven
- Scalable
The Assessment Medium

Structured
Expert-curated
Expensive

Our work

Free-form
Student-driven
Scalable

Structured
Expert-curated
Expensive
The Assessment Medium

Our work

Structured

Student-driven

Scalable
The Assessment Medium

Our work

Structured
Student-driven
Scalable
Student-generated Testing

Idea: Students both generate & answer test questions
Student-generated Testing

Value of Student Questions

Rothstein, Dan, and Luz Santana. Make Just One Change: Teach Students to Ask Their Own Questions. Harvard Education Press. 8 Story Street First Floor, Cambridge, MA 02138, 2011.


Can we Assess Implicitly?

Can students generate an accurate* assessment of their knowledge under complete autonomy?
Can we Assess Implicitly?
Can we Assess Implicitly?

Expert Ranking

Implicit Ranking
Can we Assess Implicitly?

Rank

Correlation
Student-generated Testing

Idea: Students both generate & answer test questions

Goal: 1. Discover good students
      2. Discover good questions
**Student-generated Testing**

**Idea:** Students *both* generate & answer test questions

**Goal:**
1. Student assessment
2. Test design
Student-generated Testing

Idea: Students both generate & answer test questions

Goal: 1. Student assessment

2. Test design
How to Assess Autonomously?
What ship did Darwin travel on to get to the islands?
What ship did Darwin travel on to get to the islands? 

H.M.S. Beagle
What ship did Darwin travel on to get to the islands?

H.M.S. Beagle

What birds helped to inspire Darwin's theories?
What ship did Darwin travel on to get to the islands?

- H.M.S. Beagle
- The Mayflower

What birds helped to inspire Darwin's theories?

- Finches
- Floreana Mockingbird

Which Scientific concepts were introduced by Darwin?

- Natural selection
What ship did Darwin travel on to get to the islands? H.M.S. Beagle

What birds helped to inspire Darwin's theories? Finches

Which scientific concepts were introduced by Darwin? Natural selection
What ship did Darwin travel on to get to the islands?

What birds helped to inspire Darwin's theories?

Which scientific concepts were introduced by Darwin?

H.M.S. Beagle

Natural selection

Floreana

Mockingbird

Finches
What ship did Darwin travel on to get to the islands?

H.M.S. Beagle

What birds helped to inspire Darwin's theories?

Finches, Floreana Mockingbird

Which Scientific concepts were introduced by Darwin?

Natural selection
What ship did Darwin travel on to get to the islands?

- H.M.S. Beagle

Which Scientific concepts were introduced by Darwin?

- Natural selection

What birds helped to inspire Darwin’s theories?

- Finches
- Floreana Mockingbird
- None of the Above
What ship did Darwin travel on to get to the islands?

- H.M.S. Beagle

What birds helped to inspire Darwin's theories?

- Finches
- Floreana Mockingbird

Which Scientific concepts were introduced by Darwin?

- Natural selection
- None of the above
What ship did Darwin travel on to get to the islands? 

H.M.S. Beagle

What birds helped to inspire Darwin's theories? 

Finches, Floreana Mockingbird

Which scientific concepts were introduced by Darwin? 

Natural selection

Mathematical formula:

\[ z_{ij}^l \sim \text{Bernoulli}(\theta) \]

\[ \theta = \frac{1}{1 + \exp(-s_l[a_{ij} - q_i])} \]
What ship did Darwin travel on to get to the islands?

Q

What birds helped to inspire Darwin's theories?

Q

Which scientific concepts were introduced by Darwin?

Q

H.M.S. Beagle

Natural selection

Finches

Mockingbird

The Mayflower

NONE OF THE ABOVE
Embed Students and Answers

LESS able

student

MORE able

More obviously WRONG

Q_1

Q_2

Q_3

More obviously RIGHT
Embed Students and Answers

LESS able

student

MORE able

More obviously RIGHT

More obviously WRONG

Q₁

Q₂

Q₃
Student-generated Testing

Idea: Students both generate & answer test questions

Goal: 1. Student assessment
2. Test design
Student-generated Testing

Idea: Students *both* generate & answer test questions

Goal: 1. Student assessment
      2. Test design
Optimal Test Design
Optimal Test Design
Optimal Test Design

Optimal Subset of answers
How did Darwin get to the Galapagos Islands?

Answer 1: (A Royal Navy boat)

Answer 2: (on a ship)

More obviously RIGHT

More obviously WRONG
How did Darwin get to the Galapagos Islands?

More obviously RIGHT

More obviously WRONG

\[ \mathbb{E}(\text{InfoGain}) \approx 0 \]

Answer 1
(A Royal Navy boat)

Answer 2
(on a ship)

correctness threshold
Deterministic Expected Information Gain vs. Distance between answers
415 Which Scientific concepts were introduced by Darwin?
415 Which Scientific concepts were introduced by Darwin?
Exploration vs. Exploitation (Bayesian Bandit)
Exploration vs. Exploitation

What are the three possible outcomes on the phenotype, from genetic changes caused by mutation?
What are the three possible outcomes on the phenotype, from genetic changes caused by mutation?
Evaluation

• Three baselines
• Two heuristic baselines
• Bachrach2012 (MSR)
• Metrics
• Kendall Tau
• Precision@%
• Compare to Expert Ranking
Experiment: Dataset EVOL

The graph shows the precision at different top k% of students. The legend indicates:
- Our Model
- microsoft
- baseline 1
- baseline 2
- random

The x-axis represents the top k% of students, while the y-axis shows the precision.
Experiment: Dataset SAT

The diagram shows a graph with the y-axis labeled "Precision @ K%" and the x-axis labeled "Top K% of Students". The graph compares different models:

- **Our Model**
- **Microsoft**
- **Baseline 1**
- **Baseline 2**
- **Random**

The graph illustrates the precision of different models as a function of the top K% of students.
Experiment: Evol \((x5)\)
Agreement between teachers
Expert-Expert vs Model-Expert

![Graph showing comparison between Expert-Expert and Model-Expert performance.](image-url)
The Assessment Medium
The Assessment Medium
Predicting the “winner” in a post
Predicting the “winner” in a post

Data: AskUBUNTU

110K users

Pruned to 2K
Predicting the “winner” in a post

Reputation™ 64.8 %
Upvotes (User History) 58.1 %
SVM (Rep, Upv, Age) 72.2%
Our Model 82.4%
Summary

Can students generate an accurate assessment of their knowledge under complete autonomy?
Can students generate an accurate assessment of their knowledge under complete autonomy? Probably YES*