

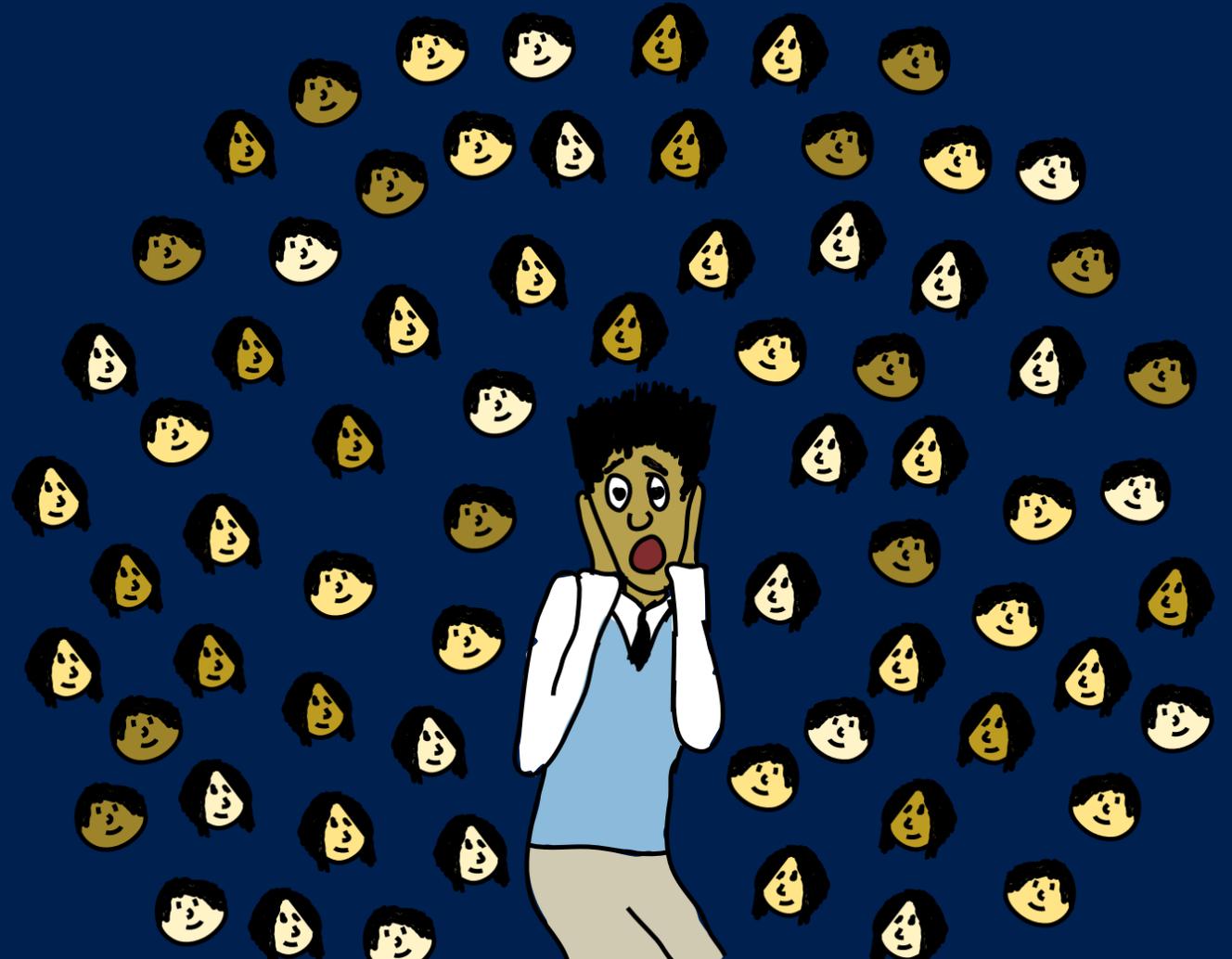


Powergrading: Helping Teachers Grade Short Answers at Scale

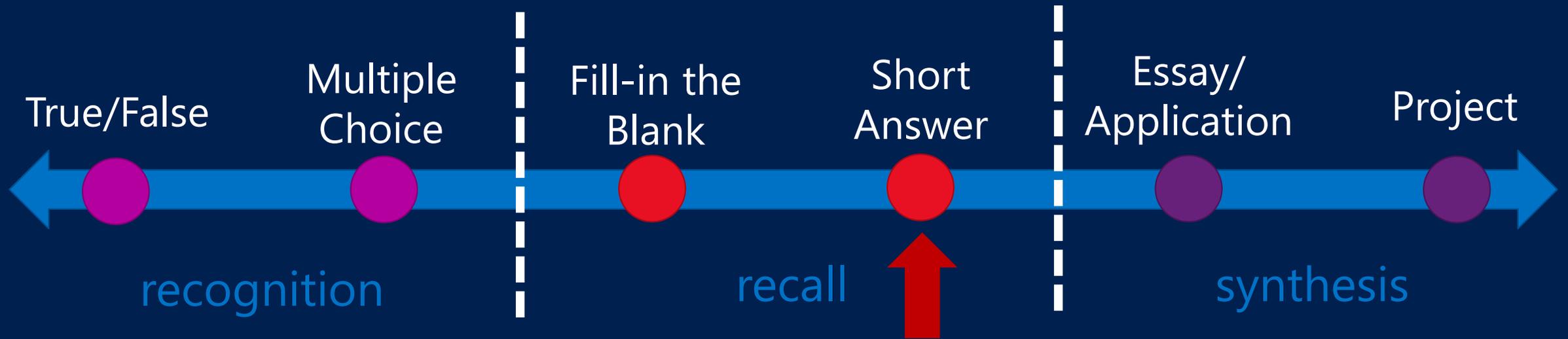
Sumit Basu, Chuck Jacobs,
Michael Brooks, Lucy Vanderwende

Challenges of Testing & Grading at Scale

- Modern web technologies are solving the **one:many** problem, but there are many challenges to the **many:one** direction...



Not all Tests are Created Equal



- Laufer and Goldstein, 2004, on the difficulty of Recall tasks vs. Recognition
- Anderson and Biddle, 1975, "On Asking People Questions About What They are Reading."

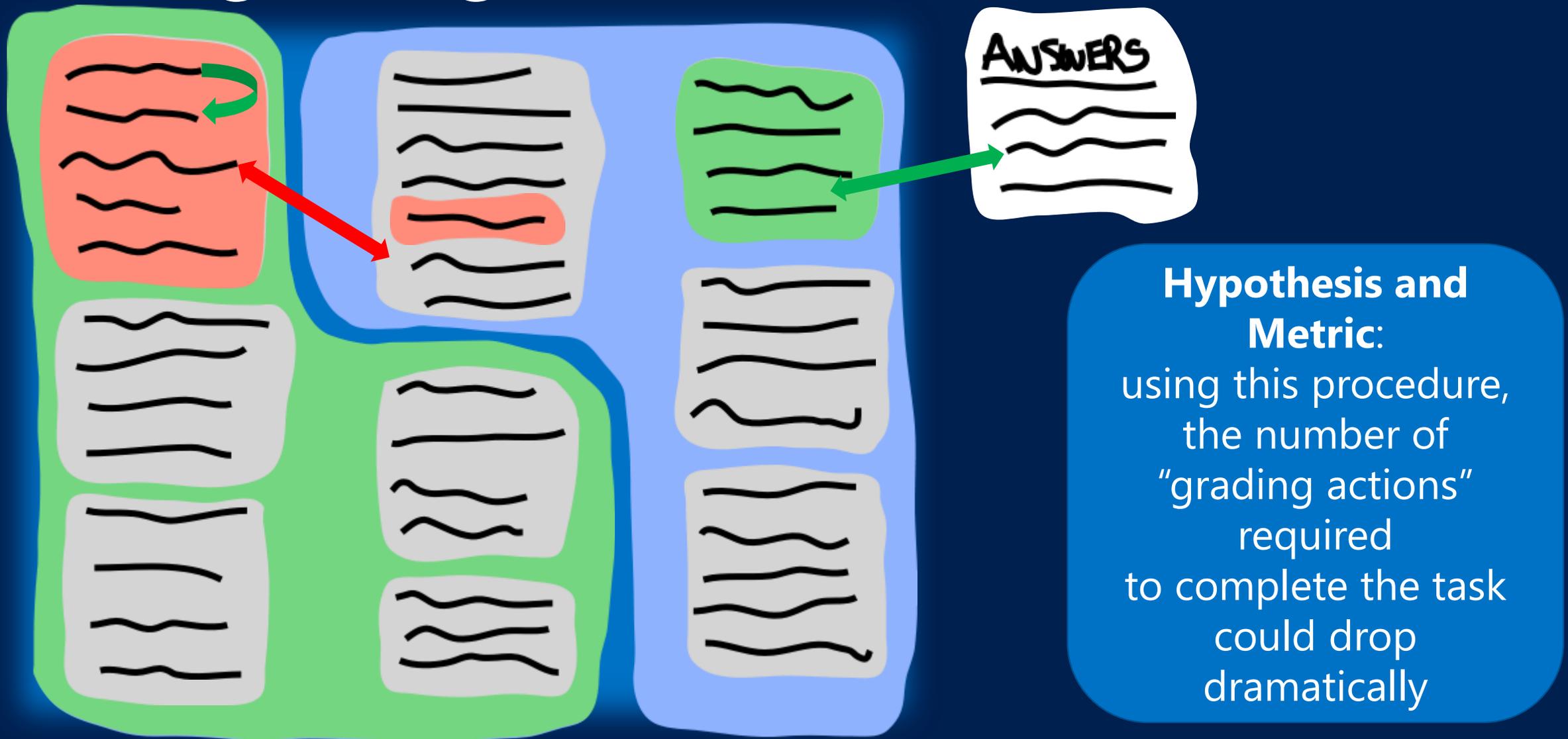
Grading is More than Assigning Scores

- Grading Practices
 - Grading is complex and individualized
 - Maintaining consistency between students is a challenge
- Instructor grading has important benefits
 - Opportunity for rich feedback
 - Getting a snapshot of student understanding
 - Adapting teaching to student needs
- Issues with Auto-Only Grading
 - some answers are always left on the table
 - no feedback for teachers
- Assisted Grading: can we leverage the best of human and machine?

Core Idea:

Perhaps clustering could help deal with a large number of answers, while still allowing for completeness, feedback, and reflection

Powergrading Overview



Actions Remaining after Auto + 3 User Actions

| | Metric Clustering | | | |
|----|-------------------|----|-----|------------|
| Q# | G1 | G2 | G3 | Num Unique |
| 1 | 2 | 1 | 2 | 57 |
| 2 | 14 | 12 | 13 | 132 |
| 3 | 80 | 87 | 145 | 586 |
| 4 | 32 | 26 | 43 | 205 |
| 5 | 20 | 18 | 26 | 138 |
| 6 | 24 | 30 | 50 | 219 |
| 7 | 16 | 12 | 11 | 395 |
| 8 | 9 | 8 | 8 | 157 |
| 13 | 75 | 54 | 50 | 367 |
| 20 | 11 | 14 | 10 | 276 |

Designing and Evaluating a User Interface for Powergrading

Clustered Grading Interface

Power Grader Question 6: Who or what makes federal (national) laws in the US? Instructions ? Done Grading

14:54 Grading Feedback

| Clusters | Subclusters | Answers |
|--|--|---|
| <p>x53 [miscellaneous]</p> <p>supreme legislature executive federal us court people laws</p> | <p>x31 [miscellaneous]</p> <p>supreme federal court us laws</p> <p>the house and senate make laws, not th...</p> | <p>x22 supreme court the house...</p> |
| <p>x27 president, congress</p> <p>legislative combination along branch passes corporations signs president presidential order can supreme veto signed sign court special effect interest groups law laws bill sent pass president's people favors senate members congress</p> | <p>x11 [miscellaneous]</p> <p>executive people</p> | <p>x2 the supreme court the house...</p> |
| <p>x16 judicial branch</p> <p>judicial branch government</p> | <p>x9 [miscellaneous]</p> <p>legislature</p> | <p>x1 supreme court. the house...</p> <p>x1 the us supreme court the house...</p> <p>x1 the supreme court makes laws in the us. the house...</p> <p>x1 suprem court the house...</p> <p>x1 suprieme court the house...</p> <p>x1 federal court the house...</p> <p>x1 supreme court? they at least verify if federal laws are constitutional the house...</p> |
| | <p>x1 [miscellaneous]</p> | |

1 subcluster selected

Mark Unread Flag

Grading

Incorrect Partial Correct

Feedback Clear Save

the house and senate make laws, not the courts

Answer Key

Q: Who or what makes federal (national) laws in the US?

Congress

Senate and House

Senate and House of Representatives



Powergrading Demo

Flat Grading Interface (Baseline)

Power Grader Question 4: What is the economic system in the United States? Instructions ? Done Grading

15:55 Grading Feedback

Answers

| | | | |
|----|--|---------------------------|-----------|
| x1 | 'mixed' economy | need capitalist or market | Incorrect |
| x1 | a bad one | capitalist or market | Incorrect |
| x1 | a capitalist economy | | Partial |
| x1 | a capitalist economy with some socialist aspects like medicare, social security, and government oversight organizations like the fda and the health department | | Partial |
| x1 | a combination of capitalism + command | | Partial |
| x1 | a corporate welfare guised as a capitalist system. | | Correct |
| x1 | a free enterprise and capitalism. | | Partial |
| x2 | a free market system | | Partial |
| x1 | a market economy | | Partial |
| x5 | a market system | | Partial |
| x1 | a market system of stocks and bonds | | Partial |
| x1 | a market system that runs the economy. | | Partial |
| x1 | a regulated free-market capitalist system | | Partial |
| x1 | a set of laws we have to follow | | Incorrect |
| x1 | a system of economics | be more specific | Incorrect |
| x1 | a total failure. | | Incorrect |

1 answer selected

Mark Unread Flag

Grading

Incorrect Partial Correct

Feedback Clear Save

Copy previous feedback:

- be more specific
- capitalist or market
- need capitalist or market

Answer Key

Q: What is the economic system in the United States?

- capitalist economy
- market economy

Questions

Preferences: Did teachers like it?

Efficiency: Was grading speed improved?

Quality: Was accuracy affected?

Feedback: Was giving feedback supported?

Reflection: Were teachers able to reflect on student answers?

Study: 25 MS/HS/College teachers, within-subjects (2 interfaces, 2 questions)

Preferences

*"When initially viewing the video on this interface, I was a little worried that it might be somewhat complicated and time consuming due to the subcategories. However, I was incorrect. This interface was **quite efficient and easy to use.**" (P15)*

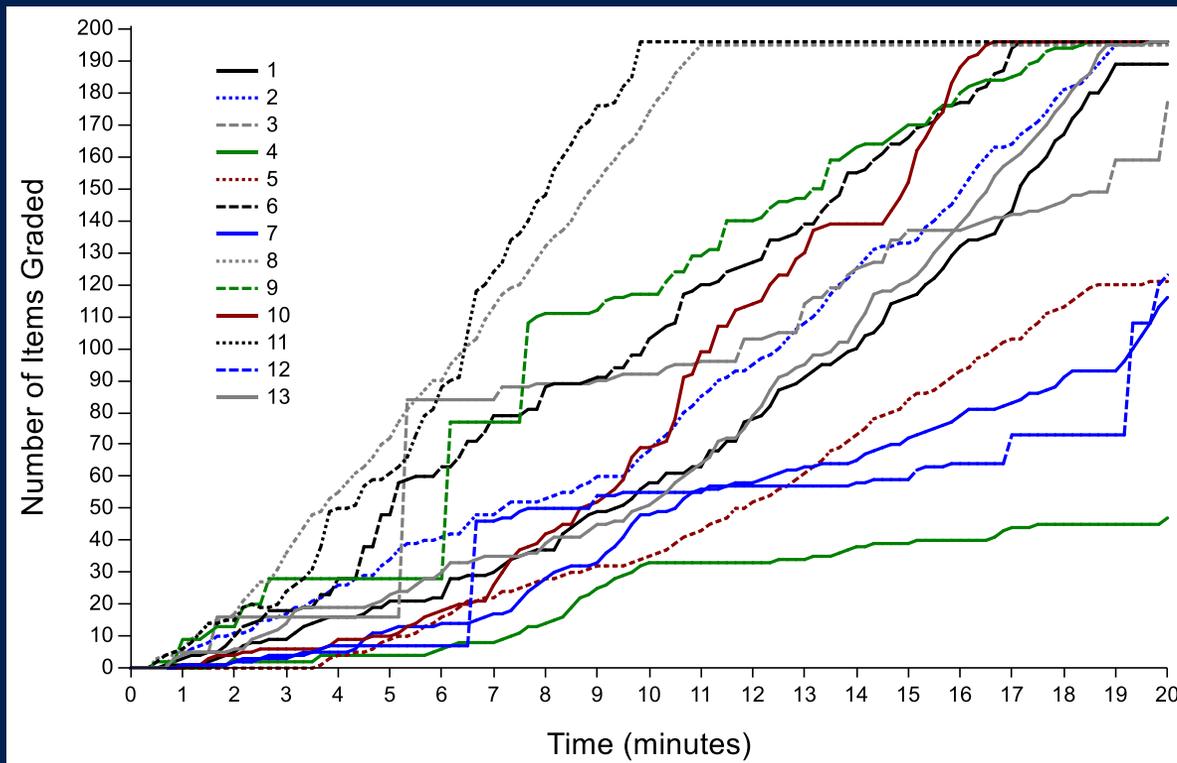
*"[The clustered interface] **worked very well** for me, especially given the large number of total responses. I found [the flat interface] quite tedious ... The clustered interface] **helped me to identify student patterns** in thinking quite well." (P12)*

| | Clustered | Flat (Baseline) |
|-----------------------|-----------|-----------------|
| Faster | 21 | 4 |
| More Enjoyable | 20 | 5 |
| Easier to Use | 20 | 5 |
| More Effective | 19 | 6 |
| Better Overall | 21 | 4 |

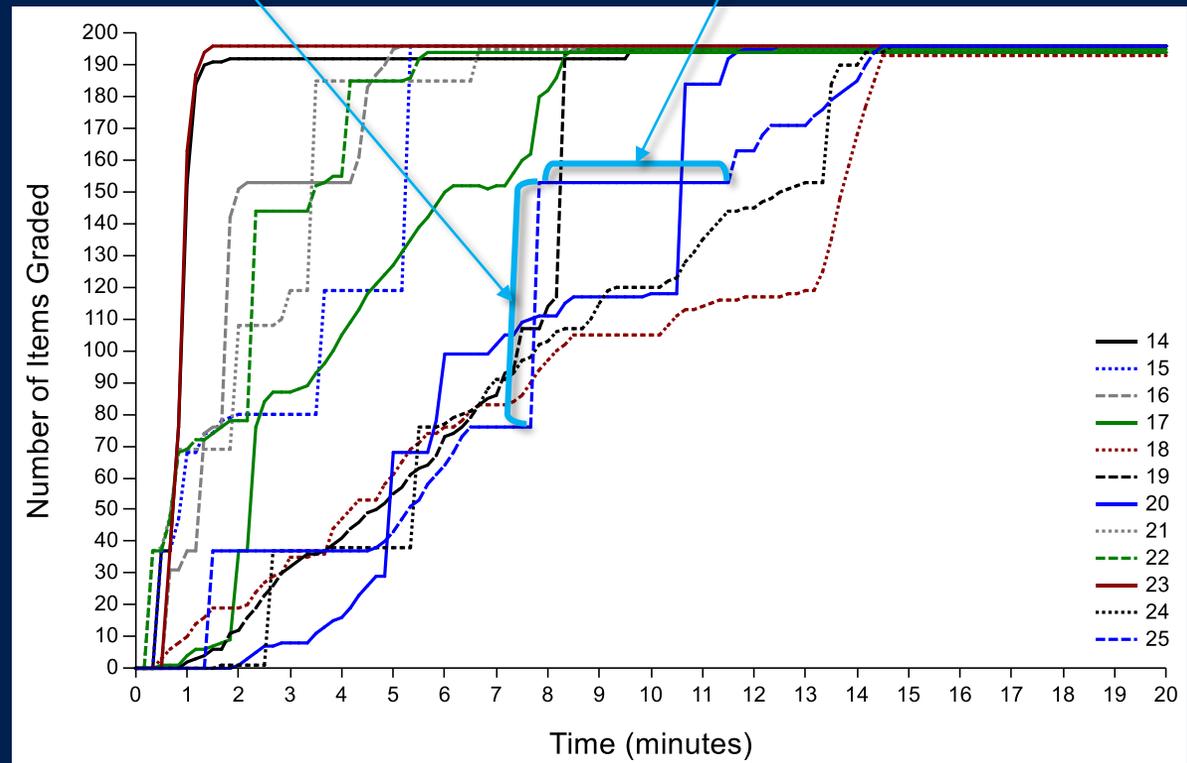
21 of 25 participants preferred the Clustered interface



Efficiency



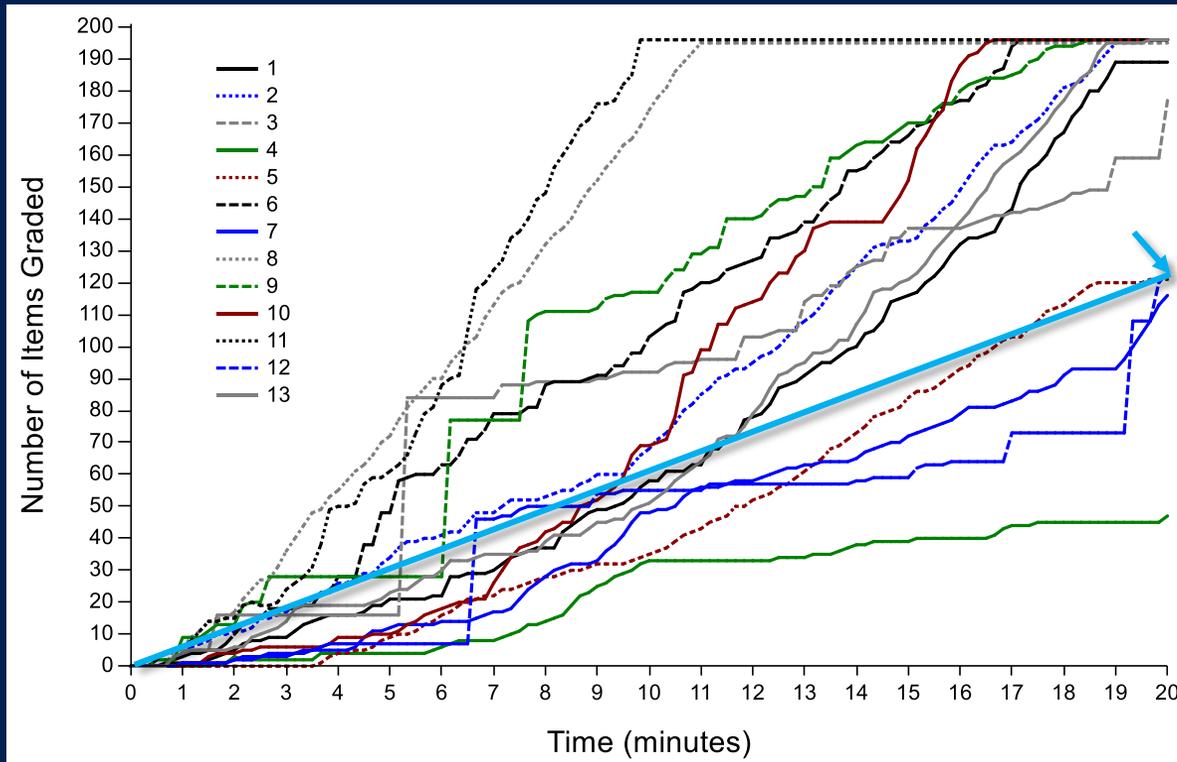
Flat Interface, answers graded vs. time for Q4



Clustered Interface, answers graded vs. time for Q4

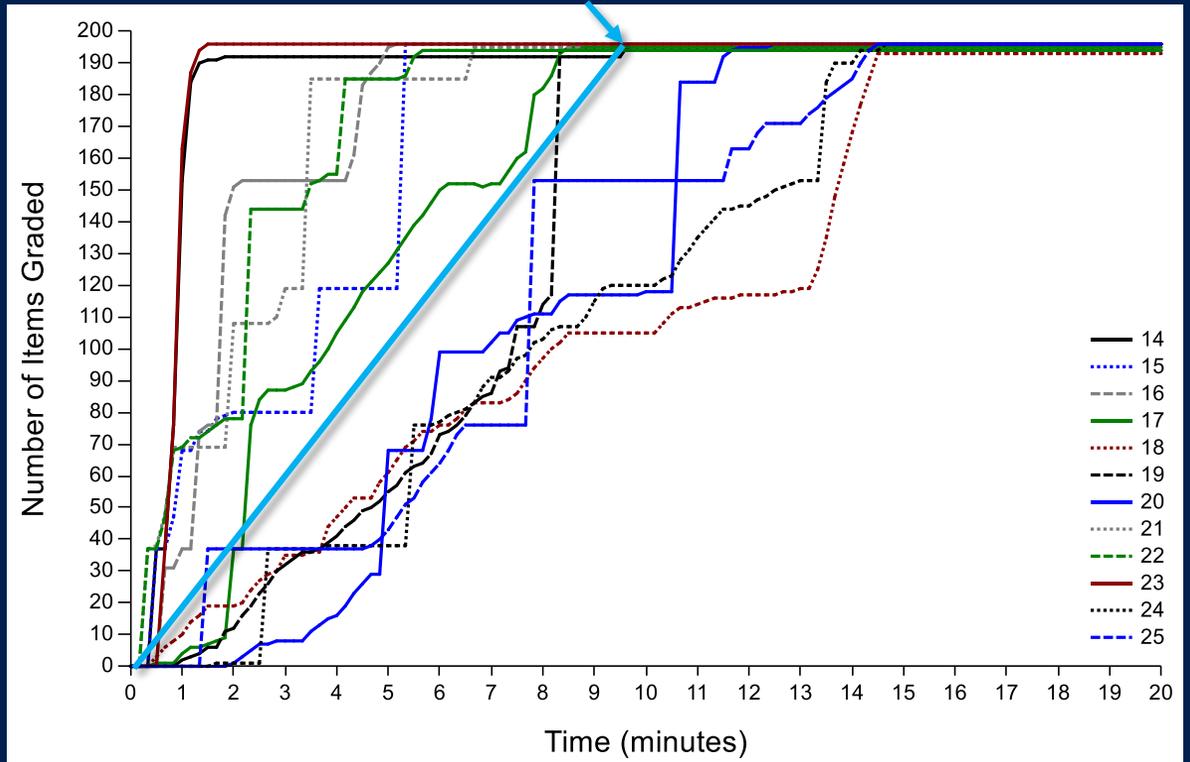
Speed Calculation

120 ans. / 20 min.
= 6 ans/min



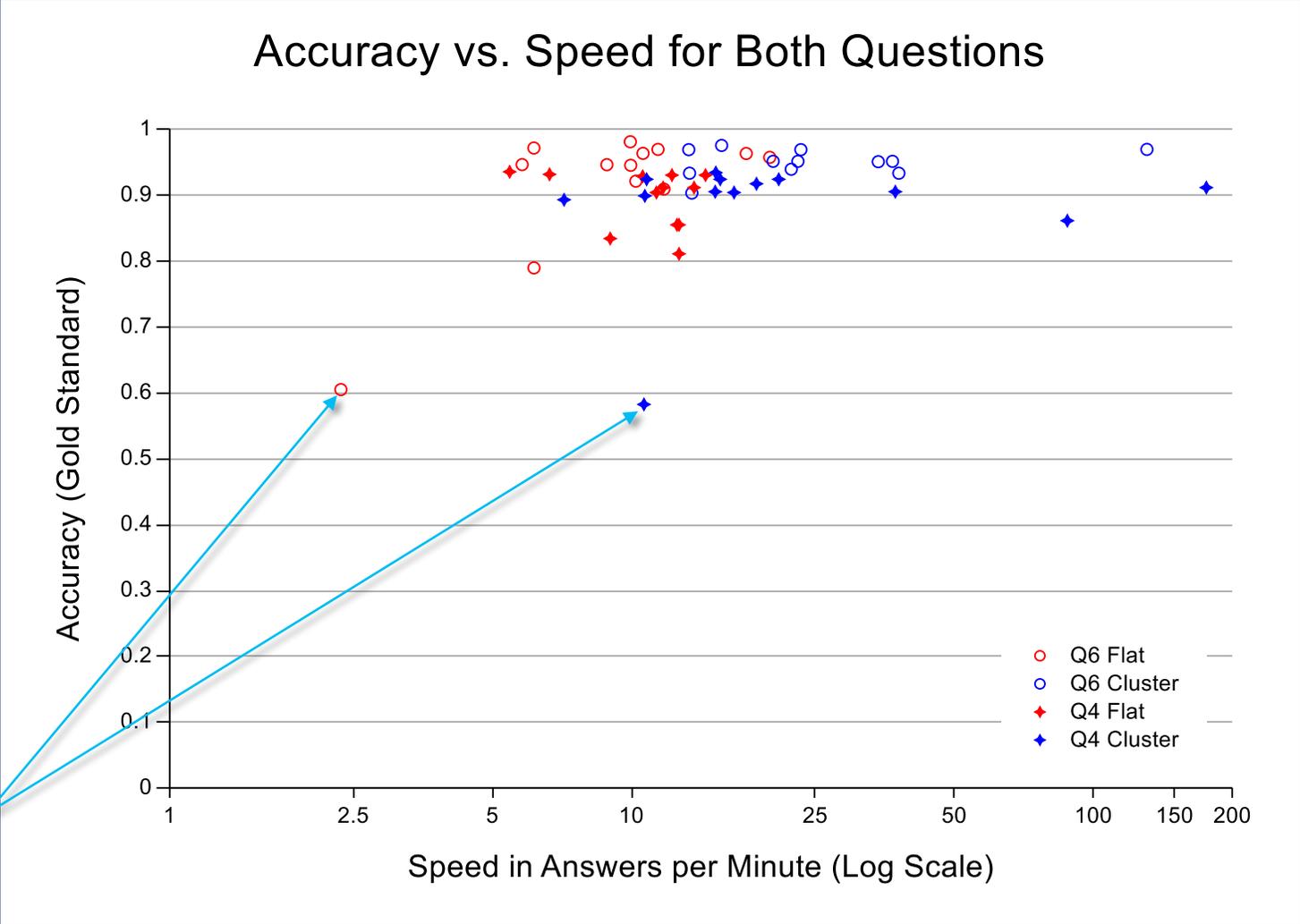
Flat Interface, answers graded vs. time for Q4

196 ans. / 9 min.
= 21 ans/min



Clustered Interface, answers graded vs. time for Q4

Grading Quality



Participant 4

Amplifying Feedback

- No difference in median instances of entered feedback
- However, clustering distributed the same number of instances to many more answers: (median 75 vs. 18)
- Clustered interface rated higher for supporting feedback

*"Being able to grade categorized responses makes it **easier on the grader and allows them to pay closer attention** to types of feedback needed." (P24)*

*"Because [the clustered interface] was so much faster, **more time could be spent giving feedback.**" (P14)*

Instructor Reflection

- Participants felt the clustered interface was better for this:

*"This interface does make **answer trends more easily identifiable.**"*
(P6)

*"I liked this [clustered] interface better; breaking the answers down into clusters allowed me to **spot patterns**, to **be more consistent** in grading, and to **devote more time to individual answers** where it wasn't clear whether they were right or wrong. The information seemed less overwhelming when presented this way, so I felt like I was less apt to mis-read or mis-grade any one answer." (P8)*

Powergrading: a “Power Tool” for Grading

- Keep teachers in the loop, but amplify their capabilities so they can handle larger classes
- Current work: making Powergrading broadly available; let us know if you want to try it out with your data

<http://bit.ly/powergrading>

sumitb@microsoft.com