

Scaling Oral Exams

Motivating Question

- How can we conduct oral exams in an efficient, fair, and cost-effective manner in a class of more than 30 students?

Why Oral Exams?

- Evaluate students independent of AI
- Mimics real world scenarios
 - Interviews, meetings, thesis defense, oral boards for physicians, quals, etc.
- Adaptable based on student performance
- Quick to grade
- Personalized feedback
- Grade by the “spirit of the law,” not the “letter of the law”
- Fun and informative

My Oral Exams

- Mimic a design interview for a tech company
- 20-25 minutes for the exam, 5 minute buffer/grading time between exams
- Identical open-ended prompts, semi-structured after
- Record the exams
- Coarse grading (✓ -, ✓ , ✓ +) along several dimensions, which collectively map to a grade
- Write down a few sentences to jog my memory
- Allow students one retake
 - Roughly 10-20% of students retake the exam

Hooray for TAs

- Up to a class size of 30ish, I can conduct all the oral exams myself
- To scale to larger classes, need to leverage TAs
 - Good practice and learning experience for TA
 - I aim for 8 students per TA
- I have historically led retakes myself

Challenges of Scaling

- Exam consistency
 - Write down the full prompt, a sample solution, and sample probing questions for the TAs
- Grading consistency
 - Training in advance, interview recordings, etc.
 - With early guidance and examples, most TA grades I double checked were well calibrated
- Room reservations (request early!)
- Feedback for students
 - TA notes + interview recordings

Grading

- Select several key topics/concepts/learning goals
- During the discussion, mentally grade them on a ✓ -, ✓, ✓+ (1-3) scale
 - Can also use - (0) or + (4) in exceptional circumstances
- If you have enough signal on one topic, guide the student to discuss another topic
 - If their knowledge seems superficial, dig deeper
- Algorithmically convert to numeric grade afterward
- If a TA is unsure about any category's grade, have them flag it for your video review

TA Cost

- TA time mostly scales with the number of students, not the number of TAs
- Assuming 8 students/TA and one exam per semester, this should take 3-4 hours for administering and grading exams and another 1 hour for prep per TA
 - This is on par with written exam or final project grading

Cost of Exams

	Written Exams	Oral Exams
Testing	Students: 1-2 hours per student TAs/Instructors: 1-2 hours per proctor	Students: 20 minutes per student TAs: 20 minutes per student
Grading	10-30 minutes per student	5 minutes per student

Student Sentiment

- Beginning of semester
 - Mix of apprehension, curiosity, and cautious excitement
 - Some anxiety, so try to reduce pressure
- End of semester
 - In course evaluations, on a 7-point scale, 83% of students responding gave a 6 or 7 (“strongly agree”) to, “*The format of the midterm (oral, intended to mimic a data engineering design interview) was valuable for my learning.*”
 - Multiple students (and TAs!) have told me that the oral exams have had a direct impact on getting an internship/job