**Research Questions:**

What were the experiences/ perspectives of the TA instructors while giving oral exams at a large scale?

What were they focusing on and talking about? In response to what?

What did we learn from attempting to implement oral exams at large scale?

I could record timing, like how long did they spend talking about certain topics in each discussion? And then proportions in survey responses as well, remembering the questions we were asking.

D1:

12:30 – number of questions (exam itself)

17:25 – about programming questions (not oral exam related)

20:45 – how to unhide the feedback (Admin, Grading)

23:35 – Anything else about the oral exams? (Admin, scheduling)

24:40 – Who makes classes? (not oral exam)

25:40 – Doing oral exams this morning again (options harder, Exam itself)

27:17 – Talking about people signing up (not related to oral exam)

28:54 - End

1. *“Doing the exam/ The administration of the exam”*

*Exam Scheduling*

D1: TA HW followed by Oral exam probably not best, leads to big instructional gap, especially with fall break, maybe can be fixed, maybe not

Try to get class not on Tuesdays, haha

Post oral exam 2 survey:

In both cases, Oral Exams have predated major exams that \*I\* sit. I do not enjoy lackluster preparation because I am quite literally drained after having to endure, most recently, 18 hours of a Zoom meeting because I had to accommodate for inevitable tech issues. This has already manifested itself in the first exam round, where I had 701 and 703 exams and homework assignments during and directly following oral exams; as for this week, I now also have two exams (558,701) directly following my oral exams. Combine this with grading, projects, homeworks, etc., and it is a lose-lose-lose for my students, my instructors, and me because it is literal attrition.

To solve these types of issues:

- Graders that jointly serve both classes, if we could get any? I know a 1-credit hour course is lower-priority for receiving gradersâ€¦ but it might make a difference!

- More autonomy in scheduling â€” it was as simple as telling students â€œdonâ€™t study vectors for the oral examâ€ this time around, i.e. modifying what wouldâ€™ve been last weekâ€™s lecture to address jointly the oral exam and new content not covered. This, seriously, would help my and other first-year instructors mental health. We were told â€œfirst-year TAs generally have lower workloads than senior TAsâ€ in an offer letter, yet it feels like that is ignored altogether in my case, whereas other instructors/first-years are not reporting the same level of issue."

"If the oral exam format goes unchanged, I am unsure if I am fit to be an instructor next semester. I will have a \*harder\* schedule in 702,704,705, and really donâ€™t see anything besides what has happened thus far in this semester happening if we donâ€™t adjust.

"

Also, having office hour days as Monday and Thursday rather than Monday and Tuesday may improve students' ability to seek help through OH as Tuesday OH tend to be dead as students have turned in assignments the night before and generally have not begun to consider next assignments yet."

D3: With HW being week before was hard, not having feedback, cause those who did not do well on HW didn’t do well on test

Hw week then Exam Week

*Student Scheduling*

Pre-Survey: The students scheduled slot are too spread out

D1: Watching out for students who are going to be late, cause it will affect other students

One person did not have a 5 min break

Another TA they did 5 min every hour

If you wait too long, some people would leave

Maybe ask students to be free 5-10 minutes after, just in case of scheduling issues

Post oral exam 1 survey:

how to make everyone register and show up on time

Post oral exam 2 survey:

"Scheduling, scheduling, scheduling.

A couple of students missed their appointments just due to misrecording when they were. A couple more missed their appointments for good reasons. These happening late in the week unfortunately means I need to try to do make-ups this week.

some people get zero points.

"If possible, avoid scheduling any time Friday so that you can make time then for any necessary make-ups and so avoid having to do make-ups the week after if someone flakes Friday.

In my experience, a few days of constant exams is better than having hours each day! :)

Also, build in more than a 5 minute break every 30 minutes - I think 10 minutes per 30 minutes is about right. Otherwise, your appointments will run over and you'll need to use your break to get back on time. (Or maybe schedule 6 minute appointments?)

Enough break time, reschedule if too late and there being following students

D3: 2-4 students make up exam

Couple missed it, need to reschedule

Wish that none were scheduled for Friday so they could reschedule later

Another student didn’t sign up for one

Student no a high achiever

One TA was scheduling Saturday time (which they liked)

People switched from Friday to Saturday

Frustrating when people don’t show up, and there needs to be a reschedule

Also the scheduling as one reason not to do it with a big class

Pre-Survey: Timeframe

Time

"1.) Students may go over time allotted.

D1:2-3 minutes behind

1/3 me, 2/3 zoom with timing issues

Handling a bathroom break (how to schedule or position yourself for bathroom breaks)

Post oral exam-survey 1:

could be improved with slightly longer time periods (so if TA builds in more breaks)

D2: One TA would send out something every 30 minutes reminding students that they were behind

*Taking care of self during administration*

Pre-Survey: the ability to stay upbeat enough to administer a large number of them back-to-back-to-back.

Interactive

D1: Handling a bathroom break (how to schedule or position yourself for bathroom breaks)

Going before (long blocks if possible)

Position yourself near a bathroom

Make yourself behind

Hey, student, could you move so I can go to bathroom?

Enable guard time, so people can’t just show up when you have an open time

Post oral exam 2 survey:

It's incredibly grueling to administer 140 oral exams, even spaced well across multiple days. Even though it takes about as much time as grading homework, you are fully occupied for that whole time - you can't snack or use the restroom or walk around if you're getting sore from sitting.

D3: like being able to get up and walk around, versus zoom the whoel time

*Navigating the use of Zoom for the Oral Exam*

D1: Talking about recording the oral exams and how that works

Talking about how much space it takes and whether it works

Other people with non-break out room works for Cloud

Could upload to google drive to save space, but didn’t take too much

On and off with the recordings from one, but most others left the recordings on (maybe every break though).

Setting it so there is a waiting room

Post Oral exam 1 survey: I think of myself as technologically literate but the oral exams have challenged that conception"

*Dealing with Technical Difficulties (moodle and zoom)*

Pre-Survey: internet issues

D1: Zoom crashing

Some bad wifi

15 minutes behind

Some people sneaking into rooms due to the mix up of links on Moodle somehow

Updated link, sent out email to let students know

Expect students to have technical difficulties, and having to do more time later in the day

More unhappy with campus wifi (other Tas agreed)

Trying to unhide oral feedback in moodle (continued discussion)

Post oral exam 1 survey:

Zoom failing to do waiting rooms mostly

D2: issues with zoom waiting rooms for one TA (already knew about this one Justin)

*Navigating Student Accommodations*

D2: One student asked if they could do the exam in person, ESL, by zoom it can be hard to hear the question clearly

Concerns about fairness equitability across students, maybe posting comments

One TA actually did one in his office hour also with DRO stuff

Keeping consistency

Post oral exam 2 survey:

Students who speak English as a second language occasionally needed a bit more time - I think it should be explicitly allowed that they can have more time, since this isn't intended to be an English test.

*Time commitment*

D1: In response to question about it, 12 hour estimate, wasn’t ridiculous, wasn’t like I was 8 hours of emails

Post Oral Exam 1 Survey:

Took way too long

Time concerns mostly. 15 hours on Zoom will drive me insane.

D2: One felt like they were losing their sanity being on zoom for so long

Post oral exam 2 survey:

"Extremely time inefficient for TAs.

overall, very time consuming.

It's incredibly grueling to administer 140 oral exams, even spaced well across multiple days. Even though it takes about as much time as grading homework, you are fully occupied for that whole time - you can't snack or use the restroom or walk around if you're getting sore from sitting.

D3:You have to schedule the time

If it were just me with 120 people, I would not do this to myself

*Grading/Technology issues*

Pre-Survey: Forgetting something someone says, difficulty with asking "follow-ups" to try and gear/pry the right answer.

The criteria could be too flexible.

Bias based on familiarity with students

Poor memory

"1.) Student answers may not fit easily into specific bins for grading.

1b.) How to differentiate between a 2-3-4 pt answer?"

D1: We have feedback section, but is there something else we’re supposed to do?

Put numbers in comments first, then clicking on categories

Clicking notify student

There was a little discussion on how to give feedback/grade through Moodle, helping one TA

D2: Rounding up some people based on whether they thought is was fair

Post oral exam 2 survey:

A study of Israeli parole hearings found parole was much less likely to be granted if the hearing was just before lunch. Avoid bias caused by time-of-day if possible with snacks or clever scheduling and you will find it easier to be fair across exams."

*Environment*

Pre-Survey: Keeping roommates at bay/background noise control

*Feelings about giving the exam*

D1: It went fine

Post Oral exam 1 survey: moodle, zoom all goes well

Post oral exam 1 survey:

Around 8 average smoothness out of 10

Very smooth

Moodle and zoom went well

Post oral exam 1 survey:

It made me very sad.

I had fun! Not sure if my students had fun.

D2: Fun, but grades

Post oral exam 2 survey:

Causes TAs (me) undue suffering and misery

I enjoyed getting to talk to my students, and I feel like it's made us more comfortable with each other. I think they're also getting a little more comfortable speaking in programming lingo. I also thought it was a very fast and effective way to see who knows what's going on in the class and who isn't with us at all.

It's frustrating when a student comes in fully unprepared

D3: It was good

*Asking questions*

D1: Prompt with follow up? (Rows or Variables)

Justin – Try to be consistent with follow up questions

D2: Follow up prompts had to be given, maybe developing better prompts based on what students may say, being consistent

Another person asked the question by breaking it into smaller questions to help scaffold the students

Asked follow up, what two types of libraries

Someone else asked what kind, made more a lot of different interpretations

Post Oral exam 2 survey:

Don't prompt too much, if student doesn't have a clear answer then move on (to save mental anguish)

*Preparing students for the oral exam*

Post oral exam 2 survey:

I would encourage future TAs to get their students talking about R or SAS first thing in the semester and enforce the vocabulary. So many of my students are averse to using the correct terminology - some are still saying "variable" when they mean "observation", so it's hard to figure out what they mean.

D3: Making first one weighted less to prepare for the second

Perhaps having ICAs have an oral component

Getting them used to practicing explaining topics, maybe groups?

R markdown presentation, for business people another TA idea

Might be too complicated

*Interpreting student responses*

D2:

More open-ended questions students meandered a bit more

Also someone else said people were lingering on a particular kind of question

Post oral exam 2 survey:

Figuring out what students are saying is like squeezing water out of a stone."

1. *“The exam itself”*

The fairness of the test (scales, points)

*Student results*

D1: Overall average and distribution? Based on question difference, and making sure it didn’t’ make crazy distribution (fairness)

D2: Looking at overall average was 10 out of 12

Someone else mentioned 85% (from discussion of grading scale, fairness)

D3: One TA had their people switch days, and scores were worse

Another said their group did much better this time

What was your average? Asking about another question that lots of people seemed to struggle with

Send individual form for student feedback

*Grading scale (fairness)*

Pre-Survey: 2.) Students may disagree with judgment about grade, and if it's a bit subjective, how do we make them feel like that's fair?"

Post oral exam 1 survey:

Overall, yes (fair grading) -- the "harsh" penalties are avoidable if students prepare

I think grading was mostly fair, though I perhaps was also more lenient

Maybe not. Some behaviors of different levels have to be divided into the same level for the rubric. Maybe we can add more levels?

D2: One TA wanted a larger scale, 10 to 20 points

Wants to distinguish between prompting and no prompting, and between those who could answer

If its clear enough, a more refined scale should happen

Looking at overall average was 10 out of 12

Maybe 20 point scale would be great (other TA)

*Questions (wasn’t related to fairness, just improving the test)*

D1: What kind of R object is blank? Maybe question was vague?

Data structure vs Data type?

Discussion about what worked, responses worked for either, etc.

I still had 5 questions on mine (there should have been 4)

Another TA had 4, so not sure what happened

Some options felt a lot harder than others for one TA. Question 2. Almost everyone got that question wrong (you should say rows)

Post oral exam 1 survey:

as the questions were relatively straightforward.

Maybe not clear what answer is expected. Questions are too broad compared to the expected answer.

D2: Maybe questions were a bit too broad for some things? Maybe some students would get them?

What is the purpose of R markdown.

More open-ended questions students meandered a bit more

Also someone else said people were lingering on a particular kind of question

D3: Changing grade on 6 point question, rescaling grades

*Comparison to other assessments (HW, tests)*

Post oral exam 1 survey:

feel like a paper quiz can serve the same goal

Post oral exam 2 Survey:

50/50 It’s complicated and disagree that it’s the same or could accomplish the same things

You could certainly \*make\* the Oral Exam a Moodle quiz, but you lose the integrity that comes with a closed-door, 1:1 environment.

If our goal is to evaluate how well students can produce knowledge "on command" then Moodle quizzes with full access to notes and sufficient time don't well encapsulate that. A Moodle quiz with questions requiring longer answers than just regurgitation of use-cases may be able to capture that to some degree but may end up measuring something different instead.

oral exam give harder questions. but if we can grade moodle quiz by human I guess they are the same. then the only difference is close book or not.

The oral exam tests whether you're able to explain the topics we cover in class, which a quiz doesn't do. \*Maybe\* if a quiz were free response, this would be similar to an oral exam, but less lenient than an oral exam - in an oral exam, if a student says something confusing, they can still get points for correcting course.

More follow-up things, easier to know how deep students understand questions

D3: “Would ya’ll do Oral exams in your own class”?

Much harder than HW, takes more time, since it has to be specific time

It’s much more like a HW or Quiz, or just another version of it, so I don’t think it’s not bad

How is it grading? “I prefer HW, more flexible” again

It does track what we want with the %>%

*Usefulness of oral exam format itself*

Post oral exam 1 survey:

Average 7 out of 10 asking about its usefulness

Mostly just reinforced ideas that I was already getting from both the homeworks and seeing students in class in regard to their uptake of material

Post oral exam 2 survey:

"One student had no clue what the %>% was (answer to the tune of ""checks if the % is greater than""), so we are catching what we want to. Did fine on other questions, though.

I enjoyed getting to talk to my students, and I feel like it's made us more comfortable with each other. I think they're also getting a little more comfortable speaking in programming lingo. I also thought it was a very fast and effective way to see who knows what's going on in the class and who isn't with us at all.

*How much time for the exam*

D3: 5 minutes might not be enough

Sometimes people can do it in 2 minutes

1. *Interaction between students and Tas (this could include follow up questions,*

D1: plenty of talk about emails being sent, apologizing to students

D2: Several TAs would get emails about things being late or reminders from students

I have something at 5:30, chill

Some students asked a TA about the rubric, and where it was

Immediate feedback may be confusing for the students unlike other exams

Tas being asked to do oral exam in person

Post oral exam 2 survey:

Figuring out what students are saying is like squeezing water out of a stone."

D3: Students seemed more relaxed, did better

Other TA had the opposite experience

Had a couple people tell one TA they liked the oral exam format

Some students who get 0 points, which is frustrating

Students answering in super vague ways sometimes

Other TA agrees, what do you do?

**Anything I see now?**

Well….my initial analysis was just wrong. Lol. There were lots of positive things said, and people thinking it went smoothly. However, most negativity came from the second one, which directly conflicted with their own tests and they probably felt super stressed out.

They had great thoughts about oral exams vs other forms, which shows the need to talk about it a bit more and explain what we’re going for.

However, providing some good training, practicing, modifying the exams, preparing for the time commitment, are all important

They are also sources of valuable feedback as well during this time. Exam changed for the better because of the feedback.

Final pieces of advice centered around scheduling, doing things like eating to not let the time commitment get in the way of rater fatigue. The big question, is, is it worth it? Do we have demonstrably empirical real data that oral exams are in fact better than a well designed test or short answer version? So far, no.

Could we find solutions to the biggest problem, the time commitment. Cause training, good scheduling, and checking in with Tas, getting and responding to feedback could all work, but lots of students are lots of students. If this isn’t fixed, well….not impossible, not crazy, but may impact reliability of exam and will impact their experience, needed training, etc.

AI could help. Doing oral exams maybe once only. Maybe for only a group of students, maybe by volunteering. Maybe more help. Until then, hard to say if truly better, not sure….