**Background:**

**Research Questions:**

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

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

**Prior experience of TAs vs Professor**

***What did the TA’s center their discussions on when talking about the oral exams?***

**1.The time commitment**

*Evidence:* Discussions, Surveys, Moodle Schedules

**D1:** 6:10 - 6:35 - one TA time commitment was around 12 hours for the week

**D2:** 8:10-8:20 - TA sending out a message every 30 minutes (I am running behind!)

12:10-12:15 - “I was losing my sanity (in a joking way) being on zoom for so long”

**D3:** 19:27 - 19:56 - TA says oral exam takes more time than grading HW, since “you have to block out time for exam”

19:57 - 20:20 - however, someone else felt the opposite, and said they felt similar to HW

25:20 25:40 - another TA also felt like grading HW was more flexible in terms of time, but not everyone feels this way

**Pre-Exam-Thoughts:** 3 TA’s off the bat say “time, timeframe”.

**After 1st oral exam:** “Took way too long”

Time concerns mostly

**After 2nd oral exam:** “Having to endure 18 hours on zoom”

Told that first year TA’s generally have lower workloads, but this was ignored all together and the time spent here was much more than other TA’s

One TA had students miss due to tech and others things, so now it’s extra time for them that has to be made up

“Overall very time consuming”, “Extremely time inefficient for TAs”

Mentioned it being 140 exams “grueling” even when spaced across multiple days

**2. Problems with the schedule not going according to plan (tech, lateness),**

**But also scheduling strategies, not just not going according to plan (two themes)?**

*Evidence:* Discussions, Surveys

**D1:** 1:00-2:55 - one TA talks about how if a student comes late, it can ruin the timing for all the students, even with a 5 minute buffer built in every 30 minutes

3:20 - 3:45 - another talks about due to zoom crashing, it caused a 15 minute delay which was handled ok since it was late in the day

3:45-4:45 - another talks about being people showing up and just doing their exams because they got a wrong link, so extra people not on the list showing up and still doing their exams when other people are not on time

5:30-6:00 - teacher mentioned for them only a few earlier in the week (3 for 1.5 hour slot), and filled up on the last day

11:00-11:18 some students had accommodations that would have allowed them to go beyond the 5 minute slots, which would have interfered with the other parts of the schedule (no accommodations needed though) **(Ask Justin)**

14:00-14:20 - TA mentions expect technical difficulties, expect it to take longer

**D3:** 11:30 - 12:20 - TA talks about having 2-4 people that had to reschedule for the next week, should have kept Friday open for rescheduled but didn’t this time, regretted it

12:20-12:35 - teacher also had some people miss without mentioning it

13:15 - 13:40 - One TA allowed Saturday, then people switched over

26:10 - 26:25, 26:45-26:56 - people missing still, having to reschedule and reschedule with them

26:00 - 27:45 - whole discussion about when to open up slots, creating multiple schedules, etc.

28:15 - 18:20 - rescheduling mentioned again

**Pre-Exam-Thoughts:** “The student’s scheduled slots being too spread out”, “Internet issues”

**After 1st oral exam: “**how to make everyone register and show up on time”

**After 2nd oral exam:** One TA had strong feelings about scheduling surrounding themselves, as in the oral exams were placed around exams, other projects, etc. It cut too much into their time.

One TA had students miss due to tech and others things, so now it’s extra time for them that has to be made up

“Students getting late or even missing”

When giving future advice, lots about scheduling. Getting in enough break time, but also making people show up on time, not doing it on Friday, all at once….different pieces of advice, all around scheduling

**3. Bias and equity**

**D1:** 7:00-7:40 - ambiguity in one question for (estimated from one TA) 75% of their students. They used data structure, but another TA used data type to deal with that question.

11:30 - 12:30 - asking about grading and how to grade a certain question, wanting a particular structure for the grades to be more accommodating **(follow up with Justin)**

**D2:** 10:15 - 11:00 - TA says “I want to give people partial credit”, got right answer vs prompting vs prompting and right answer, 20 point scale

11:30-12:00 - one TA rounded up, other said they asked extra question not on there so adjusting, “as long as you are internally consistent” the teacher said

12:20 - 12:45 - Maybe the questions were a bit too broad

14:30 - 14:40 - TA piggybacked and liked the idea of a 20 point scale to not give “too low grades”

15:05-15:25 - Discussion of the question about libraries ... .teacher would preface it saying there are two types of libraries, other TA said what type of library, which gave very different answers.

16:35 - 17:40 - TA mentioned one student wanted to do it in person because they are ESL, and it’s hard for them to understand the question clearly. Teacher worried about equity, or sharing the reading question in chat.

17:42 - 18:25 - another student was sick, so TA did an oral exam and a quiz during office hours in person. Teacher leaves it up to them, enforces make it similar to keep fairness

**D3:** 14:00-14:50 - one test question was written wrong or seemed unfair, so they talked about changing the points allotted to it

15:15-15:45 - TA talking about themselves and another TA and a question taking others a long time

21:50 - 22:05 - TA talks about how giving two exams seems fairer than one, since students seemed to feel, “much more comfortable” the second time.

28:55 - 29:35 - one TA feels like 5 minutes may not be enough for everyone, hard to rate every one well (rater fatigue the teacher says)

**Pre-Exam-Thoughts:** Worry Students may find grading too subjective

Also fear about themselves not being good graders, like, forgetting or poor memory, differentiating between different points on answers, bias from familiarity. One person mentioned too flexible of criteria

**After-First-Exam:** Someone mentioned perhaps adding more levels to better distinguish between student answers?

Gets noisy/confounded with nervousness/ speaking ability

**After 2nd exam:** eat snacks and schedule well to avoid bias from your mood (Israeli parole hearings study)

Comment about ESL students need more time, since this is not an english test

**4. Negative affect (related to themselves)**

**D2:** 12:10-12:15 - “I was losing my sanity (in a joking way) being on zoom for so long”

**D3:** 28:00 - 28:35 - One said they would do it in a smaller class, but if it were larger they would “not do this to themselves”

**After 1st oral exam:** “15 hours on zoom will drive me insane”

**“**It makes me very sad”

**After second exam:** Causes me undue suffering and misery (same person as before)

“It’s frustrating when students come in fully unprepared, it’s incredibly grueling administering 140 oral exams”

**5. Positive affect (with the students)**

**D3:** 28:00 - 28:35 - One said they would do it in a smaller class, but if it were larger they would “not do this to themselves”

**After 1st oral exam:** “I had fun!”

**After 2nd exam:** “I enjoyed getting to talk to my students”

**6. Views on Oral Exam vs normal test**

**After 1st oral exam:** “Not sure what this accomplishes over HW/ moodle quiz”, “Paper quiz could serve same goal”

**After 2nd exam:** Comments about potentially losing integrity or losing a closed book environment, and the ability to evaluate the knowledge “on command”

Discussion about how a moodle quiz could be made to be more like an oral exam, focus on a quiz requiring longer answers, free response.

Also worries about grading and being able to follow up versus not being able to in oral exam (only 2 TAs)

**7. Misc**

**D1:** 8:30 - 9:00 - one TA forgot to create blocks, and asked “How could I use the bathroom”?

**D3:** 18:35-18:55 - one TA mentioned hearing complaints but also people saying, spontaneously, “I really like this”

25:55 - 26:05 - some students were getting 0’s

**Background:**

Oral exams are given through zoom

**D1 ends oral exam talk at 28:00**

**D2 oral exam talk begins 7:50, end at 18:50**

**D3 oral exam begins at 11:00 end at 30:50**

**Previous concerns before administering the exams**

Time, having a good environment to administer them, students not staying within the schedule

**Questions**

What is moodle used for? - **grading, scheduling**

How long did it take to do them on average? - **Not sure, but we have estimates from 12.5-18 hours**

How was it graded?

What was it like in the moment?

Why do the TA’s think ya’ll are doing oral exams vs doing a moodle quiz or HW?

What were the class sizes? **80 I think?**

Was the oral exam graded in real time or not? Were things recorded or not?

**Challenges**

Technology can fail. (mentioned in way more than just the places I listed)

**D2:** 14:01-14:25

**PM1** Mention tech issues as continuing concerns

Teaching the TA’s how to administer the exams and use the technology associated with it

Mixing up links and exams, etc.

Not getting people in on the time they signed up for (falling behind)

Doing it within the 20 hours

(**D1:** Someone mentioned it took 12-12.5 hours that week)

**PM1:** Respondent 2 said it took 15 hrs (drove them crazy being on zoom that long)

**D3:** Having to block out time, vs being able to do it whenever (though some like blocked out time)....

The Oral exam itself may not be perfect (like most first time tests

**D1:** 6:50 - 7:40

had to elaborate what R object meant to around 75%, other TA agrees,

25:55-26:45 one TA had almost every person get hung up on the answer to the PROC SORT question, even with follow up questions)

**D2:** 12:15-12:40 Another TA felt the questions might be to broad so the student wasn’t exactly sure how to answer them

**D2:** 11:44-11:55, 14:30-14:40 - suggestions for 20 point scale to give partial credit for partial answers or eventually correct answers to assign (is this good or bad?)

Subjectiveness of grading?? **D2: 11:44-** TA wanted to give half credit, would round up since the scale at the time didn’t allow for that. Talks of how different people would grade…..training maybe needed? **Doesn’t seem to solve another fundamental problem of grading, that the grade is not an objective measure but depends on the teacher. Depends on the questions asked, follow up, etc. Standardized between TA’s?**

**PM2:** Trying to not let food bias you….really, fairness across exams, loss of subjectivity

**D3:** They talk about some test questions being hard, not being answered….again, that’s any test you administer

**D3:** Timing in HW not giving enough time for relearning for test….too much, too little time

**D3:** Oral exam not doing what it was supposed to be doing (students still not responding, vague, not prepared, frustrated when they got 0)

Experience of TAs

Using the bathroom haha **D1:** 8:30 - 8:35

**D1:** 10:40-10:50 People joining the meetings when they see an open spot signing up in the moment, not ahead of time

**D3:** Getting a snack, standing up, walking around, having a break, etc.

Scheduling **D1:** 23:55-24:10 (wishing the exam wasn’t right right after HW, grading for TA’s, etc.)

Student Behavior **D2:** 7:25-8:00 students being terse or upset about things not happening on time

Not immediate feedback was confusing?: **D2:** 16:10-16:17 Justin says this, what does he mean?

Student accommodations:

Not being able to understand it through zoom, preferring it in person etc. **D2:** 16:45-18:00. Is it fair if all the other students are doing it through zoom?

**PM1** May be harder for those who are very nervous or not confident in their English-speaking abilities, equitable in this regard? Probably not, though not unlike exams on paper, but with added speaking components……

**PM2** Also mentioned about second language english students need some more help, and they suggest more time for them

TA’s not being sure of the purpose or what this accomplishes over a normal exam. 2 TA’s expressed this **PM1**

TA’s being worried about their own workload/schedules **PM1**

**Mentioned Solutions**

We can see that having a debriefing meeting

**D1:** 9:45-10:05 - position yourself by bathroom, ask for the student to wait or choose a different time, etc.

**D1:** 10:45-10:55 - Adding guard times (how do you do that? How does that work with different software)?

**D2:** 12:50-13:10 - Anticipating a range of student solutions, and creating better follow up questions

**D2:** 17:00-18:00 - Solution to not understanding, or being sick, or not feeling up to the oral part is using the chat feature instead, or as similar as possible in person experience, but Justin also says we can say no to in person. Address concern, give solution.

**D3:** 13:10-13:45 - Allowing for tests on Saturdays….which one TA liked, but only this one

**Solutions in general**

Need to have a team, what would we recommend? How many TA’s per student?

What about their backgrounds?

Need to give the TA’s time to do so with their already big work load

**Oral Exam Advantages**

**D1:** 6:50 - 7:40 - Ambiguity in what an R Object was was actually sidestepped by the TA being able to offer alternative explanations or clarify the question (measuring what they know versus how well they can interpret the question and then what they know. Real world is like this)

Takes less time

**Other thoughts**

With the recordings I can hear Justin really well, but the other TA’s can be hard to hear. What is the recording set up?

**Potential questions**

How long does it take them to do the exams? I feel like we don’t need an exact answer, and this could be asked in the discussion itself.

If we are trying to ask questions with the intent of helping others potentially scale up oral exams, what would they need to know?

**Post Second Oral Exams**

*Moodle quiz vs oral exam?*

They used words like integrity, closed vs open book, knowledge on command, I think indicating that they believed one difference was that it was whether they had access to other information or not, closed vs open book.

They also talked about it being similar to a free response question quiz, but two noted that they allow for follow up, being more lenient in this way and getting a better understanding of what students are actually saying.

*Positives?*

These were all different, so no common consistent positive thing. Not everyone had something positive to say.

*Negatives?*

Everyone had something negative to say, and there were commonalities.

The number one thing was the time. One person said, “Overall, very time consuming”. Another said, “Extremely time inefficient for TAs”. Another, “It’s incredibly grueling to administer 140 oral exams, even spaced well across multiple days…..you can’t snack or use the restroom or walk around.” Another one said they felt, “quite literally drained after having to endure…..18 hours of a zoom meeting…”

Another was issues caused by the students missing exams or having to reschedule due to technical issues. One person mentioned students missing due to technical issues, and because it’s later in the week you have to schedule make up exams. Another person said almost the same thing, with tech issues with students.

*TA Affect*

Throughout TAs used words like “incredibly greuling”, “literally drained”, “literal attrition”, causes the TA “undue suffering and misery”, and other things mentioned on this page.

Positive affect? Enjoying talking with the students, but that wasn’t related to them feeling good….centered on students feeling better, doing better, and getting to know them (just one TA)....

*Advice for future TAs?*

The number one theme was scheduling. Recommending how to space out the days, not scheduling on certain days, make-ups, timing, and breaks.

**Big Questions**

With the way this was scaled up, did we lose what an oral exam was trying to accomplish?

Where were the biggest concerns and challenges?

What would we need to do differently to make this work?

Are there recommendations for doing something similar (at scale?)?

What are the grades like this semester versus the previous? Did the oral exam really do better in terms of showcasing student understanding and grading that?

**Round two: Listening to recordings again**

* Goals: contextualizing not just what was said but why.
* Listening for potentially more positive things
* Also, challenges beyond time and solutions

**D1:**

“What advice would you give to others doing the oral discussion”

00:20 – talking about recording and downloading videos, and cloud working for one TA, testing it

00:50 – uploading to google drive afterwards

Student lateness…..5 minute break for each 30 min suggested…..was not done I guess

Zoom crashing, 15 min late…..maybe an hour

Maybe far behind….people sneaking in…..people waiting…

Links were mixed up

Perpetually 2-3 min behind, but not too bad, no one too mad

(suggestion to add extra time slots, make them longer)

Q: “Was time commitment good (within 20 hours) or do you feel like it was ridiculous?”

Said it wasn’t ridiculous, lots of emails in some days, 12 hours within 20 hours

Q: What kind of R object was blank. TA asked

“What is R object”? – data structure vs data type

Q: “Any other thoughts about oral exam? Any questions from 307 people”?

Someone asked I didn’t delete blocks…..how do you handle a bathroom break?

My body just waited for me, or…didn’t think about it

Some people take faster times, so you can do it then

Go before you get into the stuff!

Get yourself near a bathroom

Maybe ask a student to move to a different time

Students are pretty chill

Justin added a guard time…add guard time where you can so people can’t just join

It went good

*Talking about grades*

11:50 – 12:10 Gave full marks on one quesiton (one TA)….might make grade distribution less painful…

Talk about grades……next week 12:30

Talking about averages….12:47

Differences in questions for rubric…..

Releasing grades, how to in moodle?

You should expect people to have technical difficulties

Technical details on emailing students, notifying them, grading

Q: “how did we deal with the zoom recording?”

One guy turned it on and off

No worries about individualized recordings is the advice they got. Just keep camera on

Resending script for SAS?

How to unhide oral feedback on moodle….

Logistics with grades in moodle

23:25: Q- “Anything about the oral exams you want to talk about?”

Nothing

25:15 – “Q: Anything with 307?”

From doing the oral exams this morning…

Some options felt harder than others, like question 2

Almost everyone didn’t get question 2

Q: Prompt with follow up?

Yeah, but they still got stuck

Maybe add more follow up? Try to be consistent! Make plan (from Justin)

**D2:**

They started taking the exam survey first!

Make sure you’re available (from student). Emails from students

Q: 307 who gave oral exams, anything you want to talk about, maybe paraphrase what you wrote about?

I want to give people partial credit/ 20 point scale

On 2 point questions.

Prompting vs no prompting, answer versus not

Q: Do ya’ll wish you had a more refined scale?

One TA is not too worried…

Looking at average….10 out of 12. 85%

Rounding up for people if they got half credit

Justin (as long as you are consistent)

Ryan, one, I’m losing my sanity..

But maybe questions are too broad in some case?

Follow up prompts to be consistent with, thinking about how students would answer (Justin)

Purpose of R markdown

Maybe more open ended floundered, meandered, etc.

Another TA feeling like not behind 5 minutes….

Q: Any other feedback? Ask another TA?

Maybe 20 point scale would be better

Lots of people lingering around some broad questions as well like Ryan

Justin (Getting better follow up questions)

Q: Any other issues with your classes?

Was rubric for oral exam given beforehand?

Some people saying not knowing how to be graded…Ryan

Student confusion with not getting immediate feedback

Oral exam, one student, discussion in person

What reason?

ESL

Concern about equity (Justin), maybe them reading the question, or something

Keeping uniformity

DRO example, did someone’s oral exam during office hours

Address concerns

**D3:**

Began with the survey!

Q: “Anything you guys want to bring up about the oral exams?”

2-4 students need to do make ups

“I wish I hadn’t done scheduling for Friday”

Justin 3 people missed, didn’t say anything

Scheduled Saturday Times, scores on Saturday much worse

Seemed like more people were more comfortable, even with harder content

Maybe studying harder

For one TA, they did not do better

Changing how much one question was worth

Q:0 points on 1 part b?

I asked them……maybe dropping points

Q: 308?

Second question difficult, maybe change?

Another TA, not much different the second time around, based on average grade

Can we offer class eval credit? No

Sending out stuff to get student evaluations

I had a couple people tell me they liked the oral exam format

19:10 Q: Do you think you’d do Oral exams? Outside of a class taught by me?

No

I think it takes more time than HW

Also the type of time, I have lots of time to grade, more flexible

Another TA felt the opposite

One drawback with oral exam, missing instructional time

HW before, people missing HW by not sticking around

But (justin) they’re adults

Grading both oral exam and grades at once hard…

Same stuff on HW wrong and on Test, need time for them to do the HW

People more comfortable on second exam….

Also, it would be great if the ICA’s incorporated more practice with oral exams or explanations

TA agrees it would be great practicing talking….but hard to do

Maybe adapting to show how to turn R markdown into presentations for Business students, would be hard (also practice)

Q: yuka, would you do an oral exam in future class

It’s like another version of a HW or quiz. I don’t think it’s that bad

Q: Grading time wise with HW?

This TA preferred HW, more flexible

Q: What about you? Another person?

Sure

I have some students who would get 0 points

Frustrating there, and rescheduling people not showing up

Maybe having a better structure for rescheduling?

Another TA said they would use this with a smaller class size

TA likes to be able to get up and walk around, get a snack, and not be on zoom

Easier with 40 people, 20 people. But if it were just me, 120, I would not do this to myself

Many people felt like 5 minutes was not enough, maybe change time?

Rater fatigue (hard to control consistency over time)

Some people will try to answer in a vague way to get extra credit…..some people just reading code. Follow up more details, but then they just repeat it.

Chain, it checks percentage? Are they just using computer? Oral exam lets the TAs see those issues.