*Literature Review*

**What angle are we coming from?:** We’re looking at this from a “more meaningful assessment perspective”, so coming from the assessment, testing, and measurement literature. Probably starting with assessment in general, how things are normally done, problems with how things are normally done (AI being the new one), oral exams as a solution, but then problem with oral exam being scale, this study trying to fill that gap.

**What’s the gap?:** No studies on doing oral exams at large scales. Giving a description of our attempt, challenges, and what the TA’s thought who did it.

**Limitations:** No student perspective, but not part of original research question so others could look into this more

Was unsuccessful…..replicating this, trying to make it successful?

Type of class will change results, though we’d argue TA concerns would be valid in this setup regardless of material

**Oral Exams**

***Benefits and costs of written take home and online tests (open book, pandemic style) (including plagiarism and AI) comparing oral exams***

Ohmann, P. (2019, February). An assessment of oral exams in introductory cs. In *Proceedings of the 50th ACM Technical Symposium on Computer Science Education* (pp. 613-619). -”we hoped to better assess the depth of student knowledge on exam topics (without question wording or guessing getting in the way).”

More student preparation, same level of perceived difficulty

Asklund, U., & Bendix, L. (2003). Oral vs. written evaluation of students. *Pedagogisk inspirationskonferens, Lunds Tekniska Högskola, sid*, 45-46. - Oral exams provide a better way to gauge student knowledge at all levels, starting hard then asking questions. Written exams not only constrain the topic but make it harder to gauge a student’s level of knowledge.

Wiggins, G. (1990). The Case for Authentic Assessment. ERIC Digest. - Essentially this is used to contrast written tests based on made up situations that are not very complex and are yes or no, and don’t directly test the thing we would like, to more authentic assessment. Others (like the clinical psych article and the gpt one) argue that oral exams are more authentic (though not necessarily perfectly authentic I would add).

Huxham, M., Campbell, F., & Westwood, J. (2012). Oral versus written assessments: A test of student performance and attitudes. *Assessment & Evaluation in Higher Education*, *37*(1), 125-136.

***Oral exam positives***

Sabin, M., Jin, K. H., & Smith, A. (2021, March). Oral exams in shift to remote learning. In *Proceedings of the 52nd ACM Technical Symposium on Computer Science Education* (pp. 666-672). - students had overall more positive emotions (hope, pride, enjoyment, and relief) than negative (anger, hopelessness, anxiety), overall enjoyed it, fairness

Gharibyan, H. (2005). Assessing students' knowledge: oral exams vs. written tests. *ACM SIGCSE Bulletin*, *37*(3), 143-147. - students feel like it is a more fair assessment of their knowledge

Oral exam allows overall impression as a whole a grade, versus grading individual parts

Roecker, L. (2007). Using oral examination as a technique to assess student understanding and teaching effectiveness. *Journal of Chemical Education*, *84*(10), 1663. - concept clarification, guidance, deeper assessment of knowledge, follow up, dynamic assessment, immediate feedback

*Beccaria, G. (2013). The viva voce as an authentic assessment for clinical psychology students. Australian Journal of Career Development, 22(3), 139-142. -* used as a more authentic assessment of what these students can actually do

Iannone, P., & Simpson, A. (2012). Oral assessment in mathematics: implementation and outcomes. *Teaching Mathematics and Its Applications: International Journal of the IMA*, *31*(4), 179-190 - “the students in our sample particularly focus the fourth of these issues, highlighting the ability of oral assessments to gauge understanding and the inability to get high marks in them simply from parroting answers.” Students also saw final closed book written exam as gold standard….measured the oral and written the same.

Huxham, M., Campbell, F., & Westwood, J. (2012). Oral versus written assessments: A test of student performance and attitudes. *Assessment & Evaluation in Higher Education*, *37*(1), 125-136. - they list 5 advantages, which I’ll copy from another article. Huxham et al. (2010) note five main advantages to oral assessments: they develop oral communication skills, they are authentic (as they more accurately mimic job interviews or the way people defend ideas in verbal exchanges), they can be seen to be more inclusive, they can be a powerful way of evaluating understanding and they are more difficult to cheat in

Received better grades like the chemistry one, not that different for others?

Felt like it was better, professional, more understanding, etc.

Joughin, G. (2010). *A short guide to oral assessment*. Leeds Met Press in association with University of Wollongong. - depth of knowledge, depth of questioning, reframing questions, student prep is more, academic integrity

Baghdadchi, S., Qi, H., Lubarda, M., Phan, A., & Sandoval, C. (2022, June). An exploratory study of student perceptions of oral exams in undergraduate engineering courses. In *ASEE Annual Conference proceedings*. - students report positively across all the classes they had (n=316, 6 classes, all with TA), liked the interaction, liked the learning, not much undue stress though stress for some, wanted more feedback, *longer time*

Newell, S. J. (2023). Employing the interactive oral to mitigate threats to academic integrity from ChatGPT. *Scholarship of Teaching and Learning in Psychology -* whole article about how orals can help with ChatGPT and how to do them

*Ramella, D. (2019).* Oral exams: A deeply neglected tool for formative assessment in chemistry. *In Active Learning in General Chemistry: Specific Interventions (pp. 79-89). American Chemical Society. -* highlights similar things ... .preparation for the real world, less cheating, potentially better for students than written tests, good experiences, etc.

***Oral Exam negatives (we’re focusing on scale)***

***Scale problem with oral exams, why written exams were used in first place***

Ohmann, P. (2019, February). An assessment of oral exams in introductory cs. In *Proceedings of the 50th ACM Technical Symposium on Computer Science Education* (pp. 613-619). - 53 students, 1 teacher, up to 30 minutes, “Scaling to larger class sizes while continuing to keep instructor time commitments reasonable and maintain consistency across students are open challenges.”

Roecker, L. (2007). Using oral examination as a technique to assess student understanding and teaching effectiveness. *Journal of Chemical Education*, *84*(10), 1663. - at most 32 students, 15 min first exam, 30 minutes second, only with struggling students

Iannone, P., & Simpson, A. (2012). Oral assessment in mathematics: implementation and outcomes. *Teaching Mathematics and Its Applications: International Journal of the IMA*, *31*(4), 179-190. - 108 students, 3 instructors, 10 minutes per student, 6 hours per instructor in one week.

Sabin, M., Jin, K. H., & Smith, A. (2021, March). Oral exams in shift to remote learning. In *Proceedings of the 52nd ACM Technical Symposium on Computer Science Education* (pp. 666-672). - 2 instructors, 30 minutes each student, 51 students total across 3 classes

Asklund, U., & Bendix, L. (2003). Oral vs. written evaluation of students. *Pedagogisk inspirationskonferens, Lunds Tekniska Högskola, sid*, 45-46. -as a discussion question they ask, “Does it scale the same way with written exams”?

Providing an Oral Examination as an Authentic Assessment in a Large Section, Undergraduate Diversity Class

***This study and the gap it solves***

*Discussion*

Bias and equity, reliability, and validity:

Asklund, U., & Bendix, L. (2003). Oral vs. written evaluation of students. *Pedagogisk inspirationskonferens, Lunds Tekniska Högskola, sid*, 45-46. - mention evaluator bias as one drawback

Memon, M. A., Joughin, G. R., & Memon, B. (2010). Oral assessment and postgraduate medical examinations: establishing conditions for validity, reliability and fairness. *Advances in health sciences education*, *15*, 277-289. - Reliability looks like it needs longer time, more topics…..only 5 minutes probably not sufficient, but then oral exams at scale are not possible without substantial resources (**look some more)**

Joughin, G. (2010). *A short guide to oral assessment*. Leeds Met Press in association with University of Wollongong. - depth of knowledge, depth of questioning, reframing questions, student prep is more, academic integrity - reliability and bias are issues, where different instructors can give different scores, also non-anonymous so biases can come from there

Gerlt, M., von Platten, J., Klöckner, M., Larsson, V., & Cedervall, T. (2023). Reducing the Impact of Bias in Oral Assessments. In *LTH: s 12: e Pedagogiska Inspirationskonferens, 7 December 2023*. - Other biases they mention are stereotype biases (based on preconceptions of groups, could be mitigated by having a mixing of people) halo biases (seeing someone as good or worse, so grading your impression of them and not their performance, based on previous experience)

Makes me wonder, then, if what we have shown is simply we did the best we could with the resources we had, and it didn’t work.

**Classroom sizes used in other studies**

Huxham, M., Campbell, F., & Westwood, J. (2012). Oral versus written assessments: A test of student performance and attitudes. *Assessment & Evaluation in Higher Education*, *37*(1), 125-136.

*The largest group they had for oral exams was 45 students, and they had 10 volunteer interviewers. So 1 to 4.5.*

Iannone, P., & Simpson, A. (2012). Oral assessment in mathematics: implementation and outcomes. *Teaching Mathematics and Its Applications: International Journal of the IMA*, *31*(4), 179-190.

*108 students, 4 people, so….1 to 30 ish.*

Roecker, L. (2007). Using oral examination as a technique to assess student understanding and teaching effectiveness. *Journal of Chemical Education*, *84*(10), 1663.

1 to 32

Kang, D., Goico, S., Ghanbari, S., Bennallack, K., Pontes, T., O’Brien, D., & Hargis, J. (2022). Providing an oral examination as an authentic assessment in a large section, undergraduate diversity class. *International Journal for the Scholarship of Teaching and Learning*, *13*(2).

*1 to 36*

Theobold, A. S. (2021). Oral exams: A more meaningful assessment of students’ understanding. *Journal of Statistics and Data Science Education*, *29*(2), 156-159.

*1 to 60*

**Time per exam per student**

Huxham, M., Campbell, F., & Westwood, J. (2012). Oral versus written assessments: A test of student performance and attitudes. *Assessment & Evaluation in Higher Education*, *37*(1), 125-136.

*15 minutes*

Beccaria, G. (2013). The viva voce as an authentic assessment for clinical psychology students. *Australian Journal of Career Development*, *22*(3), 139-142.

9-18 minutes

Iannone, P., & Simpson, A. (2012). Oral assessment in mathematics: implementation and outcomes. *Teaching Mathematics and Its Applications: International Journal of the IMA*, *31*(4), 179-190.

*10 min*

Roecker, L. (2007). Using oral examination as a technique to assess student understanding and teaching effectiveness. *Journal of Chemical Education*, *84*(10), 1663.

*15,30 min*

Ramella, D. (2019). Oral exams: A deeply neglected tool for formative assessment in chemistry. In *Active Learning in General Chemistry: Specific Interventions* (pp. 79-89). American Chemical Society.

*30 min*

Theobold, A. S. (2021). Oral exams: A more meaningful assessment of students’ understanding. *Journal of Statistics and Data Science Education*, *29*(2), 156-159.

*10 min*

**Ways others studies handled the biases from oral exams**

Huxham, M., Campbell, F., & Westwood, J. (2012). Oral versus written assessments: A test of student performance and attitudes. *Assessment & Evaluation in Higher Education*, *37*(1), 125-136.

*Standard interview protocol, only certain questions*

Gerlt, M., von Platten, J., Klöckner, M., Larsson, V., & Cedervall, T. (2023). Reducing the Impact of Bias in Oral Assessments. In *LTH: s 12: e Pedagogiska Inspirationskonferens, 7 december 2023*.

*Diverse grading, having examiners who aren’t teaching the class test the students, etc.*

Iannone, P., & Simpson, A. (2012). Oral assessment in mathematics: implementation and outcomes. *Teaching Mathematics and Its Applications: International Journal of the IMA*, *31*(4), 179-190.

Video recording to go over exams, check for biases, multiple evaluators

Kang, D., Goico, S., Ghanbari, S., Bennallack, K., Pontes, T., O’Brien, D., & Hargis, J. (2022). Providing an oral examination as an authentic assessment in a large section, undergraduate diversity class. *International Journal for the Scholarship of Teaching and Learning*, *13*(2).

*Training Tas, multiple Tas in classroom irr*

**History of oral exams (not really, just a citation talking about how it’s been used).**

**Size of statistics (intro type) classes**

NCSU sizes…..can be large, example is our NCSU

*Potential References*

Asklund, U., & Bendix, L. (2003). Oral vs. written evaluation of students. *Pedagogisk inspirationskonferens, Lunds Tekniska Högskola, sid*, 45-46.

Beccaria, G. (2013). The viva voce as an authentic assessment for clinical psychology students. *Australian Journal of Career Development*, *22*(3), 139-142.

Gerlt, M., von Platten, J., Klöckner, M., Larsson, V., & Cedervall, T. (2023). Reducing the Impact of Bias in Oral Assessments. In *LTH: s 12: e Pedagogiska Inspirationskonferens, 7 december 2023*.

Gharibyan, H. (2005). Assessing students' knowledge: oral exams vs. written tests. *ACM SIGCSE Bulletin*, *37*(3), 143-147.

Huxham, M., Campbell, F., & Westwood, J. (2012). Oral versus written assessments: A test of student performance and attitudes. *Assessment & Evaluation in Higher Education*, *37*(1), 125-136.

Iannone, P., & Simpson, A. (2012). Oral assessment in mathematics: implementation and outcomes. *Teaching Mathematics and Its Applications: International Journal of the IMA*, *31*(4), 179-190.

Joughin, G. (2010). *A short guide to oral assessment*. Leeds Met Press in association with University of Wollongong.

Kang, D., Goico, S., Ghanbari, S., Bennallack, K., Pontes, T., O’Brien, D., & Hargis, J. (2022). Providing an oral examination as an authentic assessment in a large section, undergraduate diversity class. *International Journal for the Scholarship of Teaching and Learning*, *13*(2).

Roecker, L. (2007). Using oral examination as a technique to assess student understanding and teaching effectiveness. *Journal of Chemical Education*, *84*(10), 1663.

Memon, M. A., Joughin, G. R., & Memon, B. (2010). Oral assessment and postgraduate medical examinations: establishing conditions for validity, reliability and fairness. *Advances in health sciences education*, *15*, 277-289.

Newell, S. J. (2023). Employing the interactive oral to mitigate threats to academic integrity from ChatGPT. *Scholarship of Teaching and Learning in Psychology*. - **look more**

Ohmann, P. (2019, February). An assessment of oral exams in introductory cs. In *Proceedings of the 50th ACM Technical Symposium on Computer Science Education* (pp. 613-619).

Ramella, D. (2019). Oral exams: A deeply neglected tool for formative assessment in chemistry. In *Active Learning in General Chemistry: Specific Interventions* (pp. 79-89). American Chemical Society.

Sabin, M., Jin, K. H., & Smith, A. (2021, March). Oral exams in shift to remote learning. In *Proceedings of the 52nd ACM Technical Symposium on Computer Science Education* (pp. 666-672).

Theobold, A. S. (2021). Oral exams: A more meaningful assessment of students’ understanding. *Journal of Statistics and Data Science Education*, *29*(2), 156-159.

Wiggins, G. (1990). The Case for Authentic Assessment. ERIC Digest.

Zhang, Y., & Wildemuth, B. M. (2009). Qualitative analysis of content. Applications of social research methods to questions in information and library science, 308(319), 1-12.

**Thoughts**

Were there biases at the end? If we look across instructors in a class relative to how the kids did, was it pretty different? Can we rule out class composition, or is it possible that the kids who self-selected certain sections really were different on average than other sections? Exchangeability.

What kind of prep did the TAs have?

The whole design of the oral exam process….from the rubric to the questions to what instructions TAs were given to prep