Student Emotional Response to Oral Assessments in Computing and Mathematics

Jeremiah Johnson*, Karen Jin†, and Mihaela Sabin‡
Applied Engineering and Sciences Department, University of New Hampshire
Manchester, New Hampshire, U.S.A.
Email: *jeremiah.johnson@unh.edu, †karen.jin@unh.edu, ‡mihaela.sabin@unh.edu

***Abstract*—The COVID-19 pandemic created a host of issues for institutions of higher education over the past year, including the issue of how to effectively assess student learning when courses are taught remotely. In this work-in-progress paper we present our experience using remote oral assessments in five introductory courses in two subject areas: computing and mathematics. We discuss our motivation for adopting this new assessment format and how to successfully implement remote oral assessments to replace traditional written final exams. We conducted a post-assessment student survey to understand how students responded to the oral exam format. The purpose of the survey was to gather feedback on students' emotions during and after the assessment. Our preliminary quantitative results show that overall students experienced more positive than negative emotions in all courses, though students responded differently in computing and mathematics courses. Students generally favored the oral format, and those who did not have previous experience with this type of assessment had a similar positive response to students who were more familiar with the format. We expect to shed more light on students' experience with the oral assessment as we will continue our research and conduct a qualitative study of the open-ended responses in the survey.**

***Index Terms*—student assessment, emotion, undergraduate, mathematics, computer science.**

I. Introduction

In the fall of 2020, four instructors at our department opted for online oral assessments to replace the traditional written finals in five courses: three computing and two mathematics courses. This switch was in response to the fully remote and synchronous instructional modality instituted since the pandemic at our university and maintained during the 2020-2021 academic year. Prior to the pandemic, all five courses were taught only in person. The novelty of the oral exams and their adoption in multiple courses and in two disciplines taught by different instructors gave us the opportunity to study how students felt about this new form of assessment.

Academic programs in medical and health sciences have had a long tradition of using oral assessments to evaluate student performance [1], [2]. The occurrence of oral assessments in STEM education has been sporadic despite promising findings regarding the merits of using this kind of evaluation in biology [3], computing [4]–[7], engineering [8]–[10] and mathematics [11]. The boom of remote learning caused by the pandemic and extensive reliance on video-conferencing to conduct synchronous online classes have created conditions for considering oral examinations as a viable option to assess student learning [12]–[14].

While motivated primarily by trying to eliminate plagiarism [15], [16], experience with this form of assessment has identified notable benefits to student learning: improving communication skills, focusing on conceptual understanding, and facilitating students' demonstration of what they have learned [3], [4], [6], [7], [10], [11]. Among the drawbacks, the most frequently cited are students' feelings of anxiety and stress associated with one-on-one student-teacher interactions [3], [7], [11]. In this work-in-progress study, we focused on examining how students felt about their oral assessment experience and what emotional reactions they had during and immediately after the exams. We were also interested in understanding how students perceived this form of assessment with regard to fairly evaluating their preparation, supporting their demonstration of what they have learned, and whether they were satisfied with their overall performance and liked this means of examination.

Negative emotional reactions such as despair, frustration, or anger have been observed when students work on assignments in introductory programming and mathematics courses. Such emotions are likely to lead to self-efficacy judgments [17], [18] and correlate with student performance [19]. Mindful of the impact of emotions on student work and performance, we became particularly interested in better understanding the relationship between oral assessments and students' emotions. This line of inquiry was motivated by challenges reported in similar studies [3]–[5], [7], [10], [11] among which students' feelings of stress and anxiety was a common theme.

Building on prior work that was limited to three computing courses involving two instructors [20], in this work-in-progress we extended the study to a larger student population as we included five courses taught by four instructors in two different disciplines. The larger scope of the study allowed us to pursue three lines of inquiry formulated in the following research questions:

- **Research Question 1:** What are students' general emotional response towards the remote oral assessment?
- **Research Question 2:** Do students from different subject courses, mathematics and computing, have similar responses?
- **Research Question 3:** Does familiarity with the oral examination format correlate to more positive emotional response?

- **Research Question 4:** How do students perceive their experience with oral examinations?

The rest of the paper is organized as follows. We briefly describe the educational context of our study. We then present preliminary results of our quantitative analysis. We conclude with immediate next steps and detail how current findings will guide our research plan.

II. Oral Assessment Experience

A. Courses

We designed and implemented remote oral assessments during the fall of 2020 in two mathematics and three computing courses.

- Calculus II - Traditional first course in integral calculus and math requirement for biotechnology and engineering technology majors at our college.
- Linear Algebra for Applications - Second or third year course in linear algebra. Meets a math requirement for both computer science and analytics and data science majors.
- Intro to Web Development - General education course and major requirement for information technology majors. No prerequisites, main languages are HTML and CSS.
- Programming Fundamentals - First course in programming and requirement for analytics and data science and computing majors. No prerequisites, main programming language is Python.
- Intro to Data Structures - Second year course for analytics and data science and computing majors. Prerequisites: Programming Fundamentals, main programming language is Python.

B. Oral Assessment Logistics

All five courses were taught synchronously online, with Zoom-enabled weekly class meetings. Traditional paper-based final exams were replaced by oral assessments conducted through individual Zoom sessions. Characteristics of the oral exams (number of students, exam duration and scheduling - in a single or multiple days, and whether the course had also an oral midterm) are compared in Table I.

TABLE I
Oral exams characteristics by course

Course Name	Number of students	Duration	Single/Multi Day Scheduling	Oral Midterm?
Calculus II	11	25 min	Multi	Yes
Linear Algebra	10	25 min	Multi	Yes
Programming Fund.	20	30 min	Single	Yes
Intro Web Devel.	36	30 min	Single	No
Intro Data Structures	9	30 min	Single	Yes

During the examination, students were required to have the video camera on. In the computing courses, students were additionally required to share their screen and make use of computing tools they used during the semester for in-class lab or out-of-class assignments and projects. The exam questions and programming tasks were communicated verbally, while also through the use of a digital whiteboard, Zoom chat tool, or a combination of these means.

C. Examination Questions Design

In the study of paper exams in college-level STEM courses, Laskowski et. al. [21] suggest that "long writing questions" stood out for their high reliability when compared with multiple-choice questions. To facilitate students' demonstration of conceptual understanding and higher-order skills, we adopted open format, free-response type of questions and engaged students in activities that focused on explaining problem solving and computational decisions. Since the oral assessment sessions were 3-4 times shorter than our traditional final exams, they included fewer assessment items and specific content knowledge required by more complex problems. In designing exam questions and activities, we considered assessing learning outcomes that denote higher-order cognitive processes of analysis and evaluation (based on Bloom's revised taxonomy [22]) and characterize deeper learning at the relational and extended abstract levels (according to Biggs' SOLO taxonomy [23]).

Not having all students in a course taking the exam at the same time in a proctored setting exposed the tension between maintaining consistency across all exam sessions while introducing sufficient variability to prevent students from sharing the exam content with their peers. This means that asking the same questions to maintain consistency comes at the cost of exam integrity. In the mathematics courses, the instructor opted to maintain consistency as fully as possible, asking identical questions of all students, but following up with additional questions from a larger pool of questions targeting topics related to the original questions. In the programming fundamentals and intro to data structures courses, the exam question was a problem statement parameterized by the data structure of both the problem input and output and by the type of input-output transformation (whether a map, filter, reduce pattern or different combinations of those patterns). In the intro to web development course, students were asked to demonstrate their understanding of HTML and CSS language constructs and apply them to create web pages with one of several specified content and design features.

D. Post-Exam Student Survey

At the conclusion of each oral assessment session, instructors asked their students to complete a brief anonymous online survey. The purpose of the survey was to gather feedback on how students felt about the experience during and after the assessment.

The survey had three sections. The first section was designed to help us understand students' emotional reactions, and was based on the Achievement Emotions Questionnaire (AEQ) developed by Reinhard Pekrun and his colleagues to measure various achievement emotions experienced by students in academic settings [24]. Our AEQ adaptation included seven emotions (from the total of nine - we excluded boredom and shame): four positive (enjoyment, hope, pride, relief) and

three negative emotions (anger, anxiety, and hopelessness). For each emotion construct, we kept only those items that measure affective and cognitive emotion elements that students experienced. The second section had four agreement scale items to collect data regarding student perceptions of the oral exam experience. Survey items in these sections used a 5-point agreement scale. The last section had two open-response questions. The first question asked students to describe what they liked and disliked about the oral exam. The second question asked about their preference for the oral format over the traditional written format of the exam.

III. Preliminary Results

Research Question 1: What are students' general emotional responses towards the remote oral assessment?

Student emotional responses to the oral exam format showed a stark contrast between positive and negative emotions. Across all courses students experienced and favored more positive than negative emotions (Figure 1).

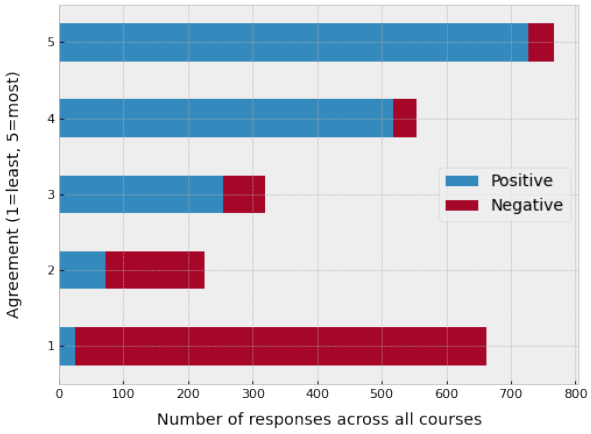

Fig. 1. Distribution of the mean response to positive vs. negative survey questions by level of agreement across all courses

More precisely, using an unpaired two-sample t-test, we found a statistically significant difference ($p < 0.01$) in the mean response to positive survey questions and the mean response to negative survey questions. Moreover, this difference was detected at the same level of statistical significance for each course individually and across all courses (Table II).

TABLE II
Mean response to positive and negative survey questions

Course Name	Mean Response (pos)	Mean Response (neg)	t-statistic	p-value
Calculus II	4.02	1.73	21.08	< 0.01
Linear Algebra	3.56	1.92	12.63	< 0.01
Programming Fund.	4.10	1.56	36.65	< 0.01
Intro Web Devel.	4.49	1.58	32.75	< 0.01
Intro Data Structures	4.28	1.27	34.31	< 0.01
Overall	**4.16**	**1.59**	**60.55**	< 0.01

We also noted a difference in the variance of responses to positive survey questions and the variance of response to negative survey questions when considered across all courses. Using a Levene test [25] for homogeneity of variances, we confirmed that the difference in variances is statistically significant ($p < 0.01$). The larger variance of 1.14 for responses to negative survey questions compared with a variance of 0.91 for responses to positive survey questions further supports the observation that overall students experienced more positive than negative emotions.

Research Question 2: Do students from different subject courses, mathematics and computing, have similar responses?

We compared the distribution of positive survey questions for the two mathematics courses with the three computing courses. Student responses in computing courses were generally more positive (Figure 2) than in mathematics courses (Figure 3).

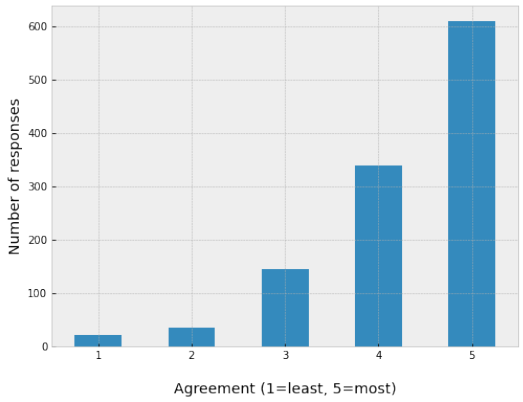

Fig. 2. Distribution of responses to positive questions in computing courses

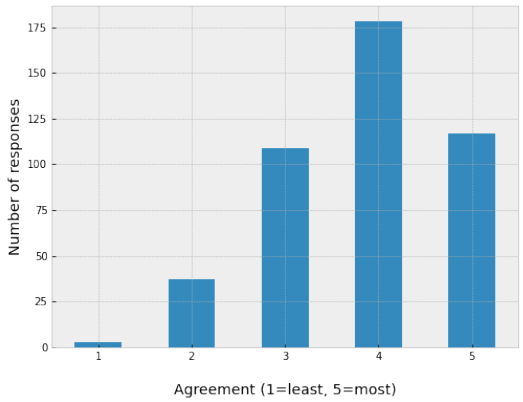

Fig. 3. Distribution of responses to positive questions in mathematics courses

Using an unpaired two-sample t-test, we found a statistically significant difference between the aggregations of all responses between the two subjects by type of question (positive or negative) as shown in Table III.

Research Question 3: Does familiarity with the oral examination format correlate to more positive emotional response? Contrary to our expectations, we found no evidence to suggest that familiarity with the oral examination format obtained by taking more than one oral assessment during the semester

TABLE III
MEAN RESPONSE TO POSITIVE AND NEGATIVE SURVEY QUESTIONS BY MATHEMATICS AND COMPUTING SUBJECT

Emotion Category	**Mathematics** Mean Response	**Computing** Mean Response	t-statistic	p-value
Positive	3.83	4.28	-8.66	< 0.01
Negative	1.81	1.51	3.70	< 0.01

led to a more positive student emotional response. The mean response to positive (resp. negative) survey questions in the introductory web design course, the only course without an oral midterm assessment, was not significantly different from the mean response to positive (resp. negative) survey questions in the other four courses in which oral midterms were administered.

Research Question 4: How do students perceive their experience with oral examinations?

As described in Section II-D, four survey questions were designed to collect data regarding overall student perceptions of the oral exam experience. These questions (with mean response and standard deviation of response shown in parentheses) were:

- Do you agree that the oral exam assesses your preparation in the course fairly? ($\mu = 4.56$, $\sigma = 0.65$)
- Do you agree that the oral exam allows you to demonstrate your [mathematical/computing] skills? ($\mu = 4.37$, $\sigma = 0.81$)
- Do you agree that you are satisfied with your performance during the oral exam? ($\mu = 4.08$, $\sigma = 0.99$)
- Do you agree that you like the oral exam as a means of assessment? ($\mu = 4.22$, $\sigma = 0.99$)

We observed that the students' highest agreement ($\mu = 4.56$) was about how fairly the oral exams assessed their preparation, followed by the degree to which this form of assessment allowed them to demonstrate their skills ($\mu = 4.37$). Although their satisfaction with their actual performance showed lower agreement ($\mu = 4.08$), students on average agreed that they liked the oral format ($\mu = 4.22$).

IV. CONCLUSION AND FUTURE WORK

We have learned from related work that oral assessments have several advantages over traditional written assessments, such as mitigating cheating and plagiarism, developing student communication skills, and facilitating close interaction between the student and the instructor. Previous research has investigated student and instructor experience with and perceptions of oral assessments and reported a common drawback: students' negative emotional responses related to feelings of anxiety and stress. Although similar studies of oral assessments in college-level computing, engineering, and other STEM programs are sparse, our experience highlights the positive responses from our students as well as some important challenges that are unique to the oral assessment format. The main finding of this work-in-progress study is that across a range of courses in computing and mathematics the student experience with final oral exams was a positive one, and, as noted in Section III, additional experience with the oral assessment format in the form of oral midterm exams did not appear to have any effect on student emotional response.

We observed some differences in students' responses between computing and mathematics subject courses. Students in computing courses had a more positive experience than students in mathematics courses. The mean positive response was the lowest and mean negative response the highest in the linear algebra course. We speculate that this may relate to the somewhat more advanced level of the course and the choice of exam problems. However, additional work is necessary to determine if the course level or the design of the exam content did indeed have an impact on student emotional response to the oral assessment.

In addition, the mean response to positive survey questions in the Intro Web Development course that did not have an oral midterm, while not statistically significantly different, was higher than the mean response to positive survey questions across all other courses. This course was the only general education course enrolled with a large number of students from all majors, which could possibly be a confounding factor. Additional work is needed to properly assess whether students' familiarity with the format impacts their emotional responses in all courses.

Our experience suggests that language barriers may present a challenge when offering oral assessments, both for students whose language of instruction is not their native language, and potentially for instructors for whom the same is true. Additionally, due to the one-on-one teacher-student interaction of the oral assessment, it is important for the teachers to recognize potential assessment bias while making intentional efforts to help relieve student anxieties that might impact their performance, for example, by providing hints and clarifications.

We also noted that although most students felt that the oral assessment format provided an accurate assessment of their knowledge, some students, particularly in mathematics, felt that the oral assessment accurately assessed their conceptual understanding, but that traditional summative assessments were still necessary to accurately assess their technical skills. This view was shared by the instructors.

Overall, the instructors felt that the oral assessments accurately evaluated the students' understanding of core concepts in computing and mathematics and met the goal of providing a more comprehensive measure of student competence. Moreover, the study results suggest oral assessments can be used effectively in both remotely-taught and face-to-face classes. We will continue our research and conduct a qualitative study of the data collected from the open-ended survey items that asked the students to describe what they liked and disliked about the oral format and whether they preferred it over a traditional written exam. We expect the qualitative themes to help us gain more insights into how to improve students' experience with oral assessments.

REFERENCES

[1] L. R. Evans, R. Ingersoll, and E. E. Smith, "The reliability, validity, and taxonomic structure of the oral examination," *Journal of medical education*, vol. 41 7, pp. 651–7, 1966.

[2] S. L. Fowell, G. Maudsley, P. Maguire, S. J. Leinster, and J. Bligh, "Student assessment in undergraduate medical education in the United Kingdom, 1998," *Medical Education*, vol. 34 Suppl 1, pp. 1–49, Sep. 2000.

[3] M. Huxham, F. Campbell, and J. Westwood, "Oral versus written assessments: a test of student performance and attitudes," *Assessment & Evaluation in Higher Education*, vol. 37, pp. 125–136, 2012.

[4] M. Dick, "Student interviews as a tool for assessment and learning in a systems analysis and design course," in *Proceedings of the 10th Annual SIGCSE Conference on Innovation and Technology in Computer Science Education*, ser. ITiCSE '05. New York, NY, USA: Association for Computing Machinery, 2005, p. 24–28. [Online]. Available: https://doi.org/10.1145/1067445.1067456

[5] H. Gharibyan, "Assessing students' knowledge: Oral exams vs. written tests," *SIGCSE Bull.*, vol. 37, no. 3, p. 143–147, Jun. 2005. [Online]. Available: https://doi.org/10.1145/1151954.1067487

[6] P. J. East and B. J. Schafer, "In-person grading: An evaluative experiment," in *Proceedings of the 36th SIGCSE Technical Symposium on Computer Science Education*, ser. SIGCSE '05. New York, NY, USA: Association for Computing Machinery, 2005, p. 378–382. [Online]. Available: https://doi.org/10.1145/1047344.1047472

[7] P. Ohmann, "An assessment of oral exams in introductory cs," in *Proceedings of the 50th ACM Technical Symposium on Computer Science Education*, ser. SIGCSE '19. New York, NY, USA: Association for Computing Machinery, 2019, p. 613–619. [Online]. Available: https://doi.org/10.1145/3287324.3287489

[8] D. B. Lowe and C. A. Scott, "Reflections on a novel approach to teaching software development," in *Proceedings of the 7th Australasian Association for Engineering Education Annual Congress and Conference*. Barton, Australia: Australasian Association for Engineering Education, 1996.

[9] D. B. Lowe, C. A. Scott, and R. Bagia, "Engineering students throw away their calculators," in *Proceedings of the 9th Australasian Association for Engineering Education Annual Congress and Conference*. Barton, Australia: Australasian Association for Engineering Education, 1998.

[10] Y. Zhao, "Impact of oral exams on a thermodynamics course performance," in *2018 ASEE Zone IV Conference*. Boulder, Colorado: ASEE Conferences, March 2018, https://peer.asee.org/29617.

[11] P. Iannone and A. Simpson, "Oral assessment in mathematics: implementation and outcomes," *Teaching Mathematics and its Applications: An International Journal of the IMA*, vol. 31, no. 4, pp. 179–190, 10 2012. [Online]. Available: https://doi.org/10.1093/teamat/hrs012

[12] D. Dumbaugh, "How to use oral examinations to revitalize online classes (Opinion) | Inside Higher Ed." [Online]. Available: https://www.insidehighered.com/advice/2020/09/09/how-use-oral-examinations-revitalize-online-classes-opinion

[13] H. Hazen and H. Hamann, "Assessing student learning in field courses using oral exams," *The Geography Teacher*, vol. 17, no. 4, pp. 130–135, 2020. [Online]. Available: https://doi.org/10.1080/19338341.2020.1828137

[14] D. Kang, S. Goico, S. Ghanbari, K. Bennallack, T. Pontes, D. O'Brien, and J. Hargis, "Providing an oral examination as an authentic assessment in a large section, undergraduate diversity class," *International Journal for the Scholarship of Teaching and Learning*, vol. 13, no. 2, May 2019. [Online]. Available: https://digitalcommons.georgiasouthern.edu/ij-sotl/vol13/iss2/10

[15] L. S. Alam, "Is plagiarism more prevalent in some forms of assessment than others," in *Proceedings of the 21st Australian Society for Computers in Tertiary Education Conference*, ser. ASCILITE'04, 2004, pp. 48–57.

[16] J. Bidgood and J. B. Merrill. (2017, May) As computer coding classes swell, so does cheating. New York Times. [Online]. Available: https://www.nytimes.com/2017/05/29/us/computer-science-cheating.html

[17] P. Kinnunen and B. Simon, "Experiencing programming assignments in cs1: The emotional toll," in *Proceedings of the Sixth International Workshop on Computing Education Research (ICER '10*, ser. ICER '10. New York, NY, USA: Association for Computing Machinery, 2010, p. 77–86. [Online]. Available: https://doi.org/10.1145/1839594.1839609

[18] ——, "CS majors' self-efficacy perceptions in CS1: Results in light of social cognitive theory," in *Proceedings of the Seventh International Workshop on Computing Education Research (ICER '11)*, ser. ICER '11. New York, NY, USA: Association for Computing Machinery, 2011, p. 19–26. [Online]. Available: https://doi.org/10.1145/2016911.2016917

[19] A. Lishinski, A. Yadav, and R. Enbody, "Students' emotional reactions to programming projects in introduction to programming: Measurement approach and influence on learning outcomes," in *Proceedings of the 2017 ACM Conference on International Computing Education Research*, ser. ICER '17. New York, NY, USA: Association for Computing Machinery, 2017, p. 30–38. [Online]. Available: https://doi.org/10.1145/3105726.3106187

[20] M. Sabin, K. H. Jin, and A. Smith, "Oral exams in shift to remote learning," in *Proceedings of the 52nd ACM Technical Symposium on Computer Science Education*, 2021, pp. 666–672.

[21] P. Laskowski, S. Karayev, and M. A. Hearst, "How do professors format exams? an analysis of question variety at scale," in *Proceedings of the Fifth Annual ACM Conference on Learning at Scale*, ser. L@S '18. New York, NY, USA: Association for Computing Machinery, 2018. [Online]. Available: https://doi.org/10.1145/3231644.3231667

[22] L. W. Anderson and D. R. Krathwohl, Eds., *A Taxonomy for Learning, Teaching, and Assessing. A Revision of Bloom's Taxonomy of Educational Objectives*, 2nd ed. New York: Allyn & Bacon, December 2001.

[23] J. B. Biggs and K. F. Collis, *Evaluating the quality of learning: the SOLO taxonomy (Structure of the observed learning outcome)*, ser. Educational psychology. New York: Academic Press, 1982.

[24] R. Pekrun, T. Goetz, A. C. Frenzel, P. Barchfeld, and R. P. Perry, "Measuring emotions in students' learning and performance: The achievement emotions questionnaire (aeq)," *Contemporary Educational Psychology*, vol. 36, no. 1, pp. 36 – 48, 2011, students' Emotions and Academic Engagement. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0361476X10000536

[25] M. B. Brown and A. B. Forsythe, "Robust tests for the equality of variances," *Journal of the American Statistical Association*, vol. 69, no. 346, pp. 364–367, 1974.

[26] M. Dick, J. Sheard, C. Bareiss, J. Carter, D. Joyce, T. Harding, and C. Laxer, "Addressing student cheating: Definitions and solutions," in *Working Group Reports from ITiCSE on Innovation and Technology in Computer Science Education*, ser. ITiCSE-WGR '02. New York, NY, USA: Association for Computing Machinery, 2002, p. 172–184. [Online]. Available: https://doi.org/10.1145/960568.783000

[27] E. Roberts, "Strategies for promoting academic integrity in CS courses," in *32nd Annual Frontiers in Education*, vol. 2, 2002, pp. F3G–F3G.