Oral Assessments as an Early Intervention Strategy

Curt Schurgers, University of California, San Diego

Curt Schurgers is an Teaching Professor in the UCSD Electrical and Computer Engineering Department. His research and teaching are focused on course redesign, active learning, and project-based learning.

Dr. Saharnaz Baghdadchi, University of California, San Diego

Saharnaz Baghdadchi is an Assistant Teaching Professor at UC San Diego. She is interested in scholarly teaching and uses active learning techniques to help students achieve an expert-like level of thinking. She guides students in bridging the gap between facts and usable knowledge to solve complex engineering problems.

Dr. Alex Phan, University of California, San Diego

Dr. Phan received his Ph.D. in Mechanical Engineering from the University of California San Diego with a specialization in medical devices. He is currently an instructor for the Department of Electrical and Computer Engineering focusing on hands-on education.

Dr. Huihui Qi, University of California, San Diego

Dr.Huihui Qi is a Teaching Assistant Professor at the University of California, San Diego.

Dr. Maziar Ghazinejad, University of California, San Diego

Maziar Ghazinejad is an assistant teaching professor in Mechanical and Aerospace Engineering Department at UC San Diego. He received his Ph.D. in mechanical engineering from UC Riverside in 2012 and holds M.S. degrees in mechanical and electrical engineering. Prior to his appointment at UCSD, he served as an assistant professor and graduate coordinator at California State University Fresno, where he received the provost faculty award in 2018. His teaching repertoire includes engineering mechanics, materials science, advanced manufacturing, and engineering design. He has also developed new classes on microanalysis, design, and nanoengineering. Ghazinejad's research on manufacturing and application of nanomaterials has produced several key journal publications, including featured cover articles, and received an Energy & Efficiency Developments (DEED) award from the American Public Power Association. He is a member of the American Society of Mechanical Engineers (ASME), Materials Research Society (MRS), American Society of Engineering Education (ASEE), and the International Society for Optics and Photonics (SPIE), where he serves as a conference chair and editor.

Dr. Minju Kim, University of California, San Diego

Minju Kim is a postdoctoral scholar at the Engaged Teaching Hub at the UCSD Teaching+Learning Commons. Minju received her Ph.D in Experimental Psychology at UC San Diego. With Engaged Teaching Hub, Minju has designed TA training materials for oral exams and have conducted quantitative analysis on the value of oral exams as early diagnostic tool (Kim et al., ASEE 2022). Minju is interested in designing assessments that can capture and motivate students' deep conceptual learning, such as oral exams and the usage of visual representations (e.g., diagrams and manual gestures).

Marko V. Lubarda, University of California, San Diego

Marko V. Lubarda is an Assistant Teaching Professor in the Department of Mechanical and Aerospace Engineering at the University of California, San Diego. He teaches mechanics, materials science, design, computational analysis, and engineering mathematics courses, and has co-authored the undergraduate textbook Intermediate Solid Mechanics (Cambridge University Press, 2020). He is dedicated to engineering pedagogy and enriching students' learning experiences through teaching innovations, curriculum design, and support of undergraduate student research.

Oral Assessments as an Early Intervention Strategy

Abstract

Oral assessments, i.e., one-on-one interview-style questioning by an instructor, have been shown to be powerful pedagogical tools. Their main benefits include the ability to assess conceptual mastery in depth due to their adaptive dialogic nature, in addition to improving students' verbal skills and serving as a tool to support academic integrity. However, assessments not only play an important role in measuring the level of students' understanding, but the assessment method also guides students' learning strategies. As such, oral assessments can serve as an important driver for students to pursue conceptual knowledge. While these dialogic assessments may exhibit challenges regarding potential bias, reliability and validity, past research has shown that with careful training and crafting these can be largely overcome. However, the main impediment towards more widescale adoption is the issue of scaling with larger class sizes, due to its reliance on one-on-one interactions between students and members of the instructional team.

To overcome the scaling issue, we have implemented and investigated an approach in which the oral assessment is only offered to a subset of students, specifically those who failed an early written exam in the course. This approach is rooted in the work on early intervention strategies. The idea is to focus on at-risk students. In this context, we do not consider the oral assessment primarily as being part of a summative assessment strategy. Instead, it is designed to be a touch point for a meaningful one-on-one interaction between a student and a member of the instructional team. The value of early interventions for at-risk students is to increase connectedness to instructional staff and resources, and student engagement and self-efficacy. The oral assessments were implemented explicitly with this focus. We also considered additional benefits, such as serving as formative assessments for the students to reflect on their level of conceptual mastery and learning strategy.

Our study was implemented in an intro to electrical engineering course, which serves as the gateway to the core curriculum. For the purposes of this intervention, the first written exam, administered at the start of week 4 of the term, was used as the tool to identify at-risk students. In this paper, we will discuss the effectiveness of this intervention, based on qualitative and quantitative data. Student feedback, obtained via a set of three surveys, shows that students find value in the oral assessment and that it increases their self-efficacy and their connectedness to the instructional team, lowering barriers to seek help. While the intervention group is too small to reach statistical significance, results hint at performance gains as well, especially when students also approach the oral assessment as a learning opportunity. These results are encouraging, as they suggest that the intervention can capture the benefits of oral assessments, while being more scalable and more targeted towards at-risk groups.

Introduction

Early Intervention Strategies

College early intervention programs, broadly speaking, are programs that try to address learning challenges for students who have fallen behind. These complement initiatives that target student preparation coming into higher education [18]. Many intervention programs are extracurricular. I.e., while students may be identified based on performance in one particular course or across courses, the intervention program itself is not part of the regular curriculum [1]. These programs come in two flavors: voluntary or required. Remedial programs that are required can enhance participation but add stigma. For this reason, participation is often elective. However, when these programs are voluntary, they rely on students' motivation, their beliefs about the usefulness and impact on themselves as learners, and free time for engagement. In addition, early intervention programs are often generic, i.e., not specific to a major. While having the potential to support a broad range of students this way, these generic programs are often avoided because students assume they will be irrelevant to their discipline [2]. As a result, they may not exhibit the desired transfer learning outcomes [3]. When extracurricular programs are targeted towards a discipline, they typically focus on connection-making and sense of community, while impact on academic performance is often inconclusive [17, 22].

Other early interventions have been integrated into STEM courses themselves, where both the student identification and intervention are entirely contained within the course. Lavelle and Keltie describe an intervention for engineering students in a calculus class [19]. At-risk students were identified based on their performance in the first written exam and then invited to meet twice with an academic advisor. The intervention was elective, leading to low participation (12%), and no statistical impact on academic performance could be concluded. Vandenbussche et al, on the other hand, incentivized participation [20]. After a first written test, students were offered a remediation path and then offered a chance to retake the test. The scores on this first written test were used to determine the remediation type. For lower performing students, this included a requirement to attend office hours. The intervention was reported to be correlated with student success, however without claiming a causal effect. The framing as remediation had some negative impact on student self-perception and behavior in the course. In another study in an engineering physics course, Bennett et al also divided students based on early written exam performance. Different intervention paths, framed as success enhancement, were then offered with participation incentives [21]. For lower scoring students, this intervention path was the opportunity to retake the earlier exam. It was noted that attending help sessions or meetings with an instructor were part of an earlier implementation of their intervention but were found to not be effective here. As with the other studies on course-integrated early intervention, academic performance was challenging to evaluate. All three studies report a correlation with academic success and anecdotal positive effects but could not claim a causal relationship [19-21].

The above-mentioned early interventions attempted to integrate aspects of student and faculty interaction. Research has shown that interactions with faculty increase student satisfaction and performance, provided they are high quality interactions rather than felt as being remedial [26-29], which may explain the mixed success in some of the early intervention programs. Rewarding faculty-student interactions have also been positively correlated to various affective outcomes among students, such as personal and social development and sense of belonging, as well as positive self-concepts and greater academic confidence [28, 30]. Related to the latter, student self-efficacy more generally, i.e., their beliefs about their academic capabilities, is an important contributor to student success. Self-efficacy is known to have a positive impact on effort, task persistence and motivation [24], and in turn on academic performance [25].

While the positive effects of attending office hours, as key facilitators of student-faculty interactions, are therefore well supported, research is mixed on the best ways to encourage office hour attendance. For example, the effect of instructor approachability is unclear. Some work suggests that it is important, supported by the fact that it is one of the main drivers as reported by students [26]. Other work claims that it has limited effect [27]. On the other hand, both sets of studies agree that perceived usefulness is an important factor [26, 27]. Students who perceive the instructor's feedback as useful are more likely to attend. Also, the act of attending office hours has a positive effect on this perception, suggesting that a less favorable perception is often due to unfamiliarity [26]. This also emphasizes the importance of breaking down that initial hesitation barrier [26].

In summary, research supports that early intervention pathways (i.e., the steps towards improvement that at-risk students are asked to take) are more successful when they promote self-efficacy, can provide an entry point into office hours, and are incentivized without attaching the stigma of remediation. In existing work, once at-risk students were identified, the intervention pathway was for them to meet with an advisor [19] or participate in another written test [20, 21]. We will evaluate whether oral assessments can serve as an effective early intervention pathway instead.

Oral Assessments

While oral assessments can refer more generally to any form of dialogic test, we specifically look at one-on-one “interrogation” of the student by a member of the instructional team [4]. These dialogic interrogation-style assessments have been shown to be able to probe the higher levels of Bloom’s taxonomy [5-16]. Research has also shown that oral assessments have several other salient benefits, beyond serving as mere tools to measure student performance. By having to explain their thought processes, students are forced to consider connections and relationships, which helps them realize where they are struggling and which concepts that they have yet to master [8]. In fact, this self-realization may also occur as they prepare for the test, by

contemplating the questions that might be asked or mimicking/practicing with friends. This anticipatory learning effect, i.e., studying differently depending on the expected nature of the test, can serve as a motivator to seek out more conceptual understanding rather than rote memorization [10, 23, 31]. Second, oral assessments have also been shown to produce affective benefits. Specifically, they can positively impact their sense of belonging/connectedness [14], which is one of the three drivers towards intrinsic motivation in the Self-Determination Theory (SDT) of motivation [32]. More intrinsic motivation leads to higher-quality learning and creativity. Also, because of the one-on-one interactions, oral assessments have been shown to increase the comfort level of students with the instructional team, which can reduce the hesitation on the part of students to visit office hours and seek help [14]. This is particularly important for members of underrepresented groups, who often have greater hesitancy to take advantage of these resources and opportunities [33]. Finally, the dialogic interactions and constructive feedback during the oral assessment can also contribute to the students' self-efficacy and feeling of competence, one of the other drivers in SDT towards intrinsic motivation [32].

These affective and academic benefits of *increased relatedness, self-efficacy and learning self-reflection* are key considerations for the adoption of oral assessments, beyond their use as pure assessment tools. In fact, these benefits are the key reasons for us to consider them as the intervention pathway in our proposed early intervention strategy. On the flip side, potential concerns that have been raised about oral assessments are increased test anxiety, consistency between evaluators and effects of implicit bias [4]. Research has shown these to be manageable through effective design and evaluator training [14]. The main challenge towards more widespread adoption of oral assessments in engineering courses has been the amount of effort involved, particularly in terms of scaling towards large class sizes [4, 6]. Only a few studies of considered course with enrollments of 100 students or more [8, 13, 16]. Their proposed solutions rely on involving the entire instructional team. Still, the amount of effort involved, including the logistics and coordination, remains one of the key challenges [14].

Methods

Design

While oral assessments have been reported to have powerful potential benefits in terms of increased relatedness, self-efficacy and learning self-reflection, the main question is whether these outweigh the cost in effort. Our observation is that these benefits are particularly important for students who are struggling in the class, who are the most at risk and have the most room to grow. This meshes directly with early intervention strategies, which specifically target low-achieving students. In fact, we believe that these two concepts align very well and complement each other naturally: oral assessments promise the affective and learning benefits that are key to

successful early intervention, whereas early intervention programs offer a vehicle to focus efforts where they are most impactful by targeting at-risk students.

> *Our contribution is to propose and evaluate a course-integrated early intervention approach where the intervention pathway for at-risk students consists of an oral assessment with a member of the instructional team.*

Below, we will describe the specifics of how we implemented this approach in one of our courses.

Implementation and Participants

Course: We integrated our proposed approach into an introduction-to-electrical-circuits course in the Electrical and Computer Engineering Department at a large public university (UC San Diego), in Fall'22. This course is the first engineering course for most students in the major and the gateway to much of the remainder of the curriculum. It has a large enrollment, with around 300 students in Fall'22 (303 students were enrolled at the start of the quarter; 279 participated in the final exam). The instructional team offered 40 hours of tutoring/office hours each week, spread throughout the week. This was done to provide as much help as possible, and to allow maximum scheduling flexibility for the students, as research has found that convenience is an important factor in whether students attend office hours [27].

Identification of at-risk students: The first written exam served to identify the at-risk students. It was administered on Monday of week 4 of the term (the term consists of 10 weeks followed by a final exam week). A similar approach was taken in related work on course-integrated early intervention strategies [19-21]. As in this related work, our identification also did not make any prior assumptions about at-risk students based on gender, ethnicity or other demographic factors. Instead, students who scored less than 6 out of 12 (including those who missed the exam) were identified as "at risk"; they were subsequently invited to take part in an oral assessment. Electing to partake involved choosing a 20-min slot from a range of offerings, which were scheduled over multiple days and did not coincide with any course meeting times (lecture, discussion, lab).

Incentive: To encourage students to take part when they were invited, the oral assessment came with the incentive of extra credit. This extra credit was up to two points, based on performance in the oral assessment, and would be added to their first written exam score, with the new score capped at 6/12. This cap was put in place to ensure that students who were offered the chance to do the oral assessment would never end up with a total exam score higher than those who were not offered this opportunity, for reasons of fairness. The first exam represented 13% of the total course grade; and it could be replaced by the final exam score if the latter was higher. As such, the extra incentive offered was minimal. Also, care was taken to frame the oral assessment as an opportunity to show their understanding, rather than as a remedial intervention, as prior work has

shown this messaging is important [20]. In this respect, framing as extra credit versus straight up credit was also beneficial.

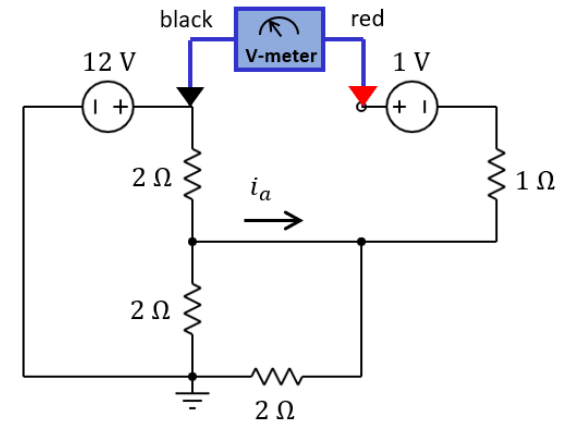
Assessors: The oral assessments were administered by the instructor and senior instructional assistants (teaching assistants and undergraduate tutors), but students did not know in advance who they would be assessed by. Each instructional assistant received training, which included a set of dedicated videos on effective communication strategies, communicating a growth mindset, the challenges of implicit bias and how to manage anxiety, reduce cognitive load and provide thoughtful feedback. This was followed by an in-person briefing with the instructor, which covered the logistics, as well as how to engage students during the assessment, help them self-assess their conceptual understanding and encourage them to attend office/tutoring hours.

Oral assessment: The oral assessments were held in-person. Two meeting rooms in the department were reserved for this purpose. Each student's oral assessment was 20 minutes. A week prior, students were given two short questions that would serve as the basis for the oral assessments. These new questions were similar in difficulty level and covered the same topics as the first written exam that had served to identify at-risk students. For this course, these questions consisted of simple circuits. They are shown in Figure 1, on the left side.

During the assessment, students were first asked to explain their approach, supplemented by probing follow-up questions about why they made certain assumptions or simplifications. The next level of questions involved asking if some alternative approaches were possible, and why or why not. Finally, the assessor would make changes to the circuits and then have the student explain how this would impact the solution and why. This process was followed for both circuits.

The instructional assistants were provided with a script which listed questions and modifications, which are shown on the right side in Figure 1. The script was given as a scaffold and guide, but assessors were encouraged to adapt to the situation with the particular student. As such, each assessment was different, and while some students got through almost all questions for each circuit, for others the focus remained on just some subparts. The main goal was to engage in a conversation with the students as soon as their edge of conceptual understanding was reached. Essentially, the succession of questions in the script were meant to probe where that edge was for each student. Throughout the process, the assessor would provide feedback and explanation, often via the Socratic method. There was also emphasis put on making sure the students felt comfortable and encouraged, while providing a vehicle to identify misconceptions. The final part of the assessment was a brief recap, with encouragement to visit tutoring/office hours.

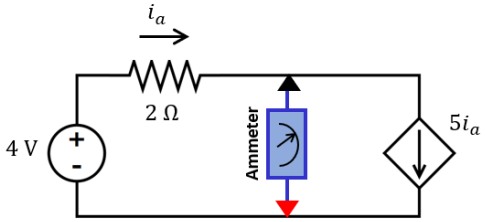
What is the current i_a? What is the reading X of the voltmeter?



You solved these problems by yourself, correct?

1. Can you explain how you approached this problem? Follow up: Why is the current through the resistor equal to zero?
2. Are any resistors in series/parallel?
3. How would things change if you replaced the 12 V voltage source with a current source?
4. What happens if you flip the direction of the 12V voltage source?

What is the reading X of the ammeter? What is the power supplied by the dependent source?



1. Can you explain how you approached this problem? Follow up: Why do you consider the ideal ammeter a short?
2. Can you do a KVL around the outer loop?
3. How would I connect a second ammeter to measure the current through the dependent source?
4. What happens if we flip the direction of i_a?
5. What is the power received by the ammeter?

Figure 1. (Left) Questions as provided to the students before the oral assessment, (Right) Assessor script used during the oral assessment.

Grading: The assessor would grade each student on a scale from 0 to 1 for each of the two circuits, resulting in a total score between 0 and 2. The grading rubric is shown in Table 1. This rubric was also shared in advance with the students.

Table 1. Oral assessment grading rubric.

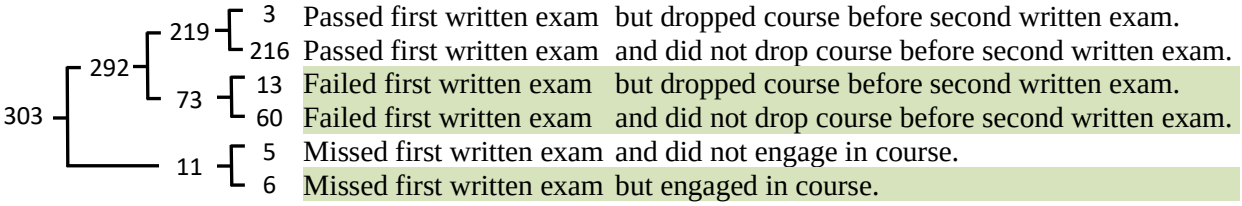
Score	Description
0	The explanation of the solution approach and/or elaboration responses have significant conceptual errors. Or the student did not show up.
0.5	The explanation of the solution approach and/or elaboration responses have some conceptual errors.
1	The explanation of the solution approach and/or elaboration responses have minor conceptual errors
1.5	The explanation of the solution approach and/or elaboration responses have virtually no conceptual errors.
2	The explanation of the solution approach and/or elaboration responses have virtually no conceptual errors. The follow-up questions regarding what happens when changes are made to the circuit are answered correctly as well.

Below are some examples of the types of conceptual errors that students made during the oral assessment. This list is not exhaustive; it includes representative examples to further elucidate the grading rubric.

Examples of minor conceptual errors:	E.g., incorrect use of passive sign convention, failure to correctly identify resistors in series/parallel, unable to answer questions after circuit modifications, etc.
Examples of medium conceptual errors:	E.g., did not realize KVL works for all loops in a circuit, not including the V-meter voltage in KVL, mixed up the internal resistor an ideal voltmeter and ammeter (and cannot explain why), etc.
Examples of major conceptual errors:	E.g., unable to perform KCL (ignoring branches), incorrect use of ohm's law (applying it on a source; using a node or source voltage and a connected resistor), etc.

Data Analysis

Participation: Out of 303 total, 292 students took the first written exam. A detailed breakdown is shown below. Of the 11 students who missed that exam, 6 did so for external reasons, such as illness or other obligations. However, another 5 did not participate in any of the course components (three written exams during the term, homework, or labs), except for the final exam. This was possible as the final exam could replace the three written exams in the total grade calculations. As such, we do not count these students as really engaging with the material, and would not expect them to take advantage of the oral assessment.



All students who failed the first written exam (i.e., scored less than 6/12) or missed it, were invited to do the oral assessment. This represents 28% of the students, which speaks to scalability. The rows marked in green are the 79 students who would have been expected to participate. Of those, 49 signed up (62%), although fewer (38%) in the subgroup that dropped the course before the second written exam (in week 7 of the term). Some were unable to attend the assessment for various reasons, and in the end 40 students participated. The reasons why some eligible students chose not to sign up or why some of those who had signed up did not attend, will be discussed in the Results section.

Data collection: Student performance on the first written exam, the oral assessment (if the student participated) and the final exam were tracked. In addition, feedback from students was solicited via three surveys.

(1) Pre-survey (189 responses received): administered after the students learned their score on the first written exam (that served to identify at-risk students) but before the oral assessment (there was one week in between to allow for the grading of the written exam and give time for students to sign up for the oral assessment).
(2) Post-assessment survey (215 responses received): administered a few days after the conclusion of the oral assessments.
(3) An end-of-term survey (151 responses received): administered during the last week of the term (before the start of finals).

These surveys were sent to all students in the class. Some survey questions were customized depending on whether students were eligible for/took the oral assessment or not. Surveys were not anonymous. However, they were managed by a member of the research team who was not the instructor in the course, and the instructor did not have direct access to the survey results. Only de-identified data was shared with the instructor.

Results

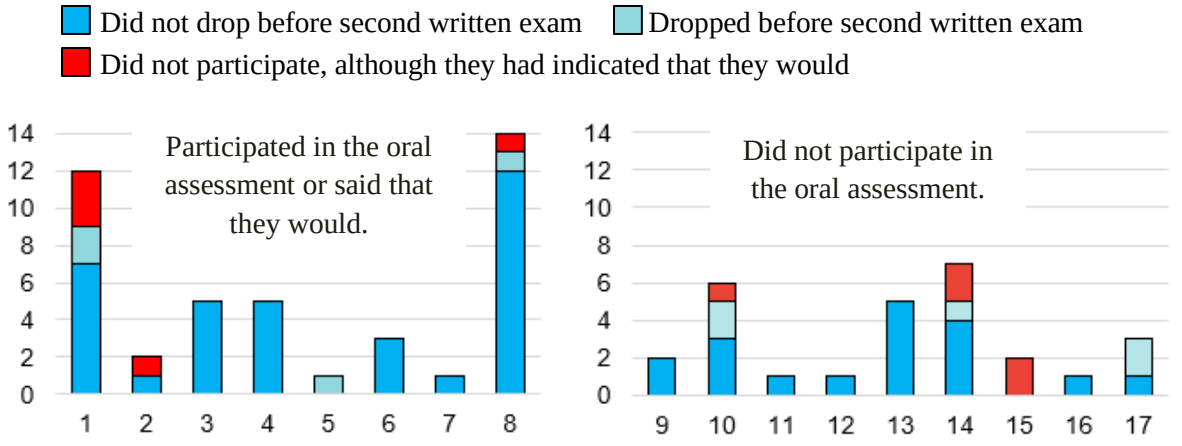
We will investigate the affective and performance-related impacts of our intervention, specifically as they relate to increased relatedness, self-efficacy and learning self-reflection.

Learning Self-Reflection and Self-Efficacy

The first aspect that we investigated was whether the intervention had a positive effect on how students approached their learning and their motivation to learn. First, we wanted to know why some students who qualified for the oral assessment decided to participate, while others did not. A free-response question to this effect was added to both the pre-survey and post-assessment survey. The answers were subsequently coded into categories, which are shown in the tables in Figure 2. The two surveys were combined to capture students who responded to both surveys, or to only one of them. Figure 2 shows the students who participated in the oral assessment in the left figure and those who chose not to in the right one. The students corresponding to the red label ("Did not participate, although they had indicated that they would") appear in both figures; these are students who on the pre-survey indicated that they would participate (with their reasons marked in the left figure) but then ended up not doing so (with their reason marked in the right figure). The other students (blue labels) were split into two subcategories: whether they dropped the course before the second written exam or not.

While most students report to be motivated by the learning benefits of the oral assessment, a significant number participated because of the extra credit. As such, we believe that providing this extra credit as an incentive is valuable. On the other hand, when looking at the reasons why some eligible students decided not to participate, we notice that some did not feel ready. This

hints at the need to promote the intervention as both an "assessment for credit" (to encourage the students who enrolled because of it) and as a valuable learning opportunity regardless of their preparation. In retrospect, it would have been worthwhile to make the latter point clearer. This might also have motivated students who decided not to participate because they missed the written exam (#17). While oral assessment slots were spread out over a full week, scheduling conflicts (#13) or external reasons (#14) may be hard to avoid; although in those cases, emphasizing the learning benefits and extending an open invitation to talk through the same problems in office hours is valuable.



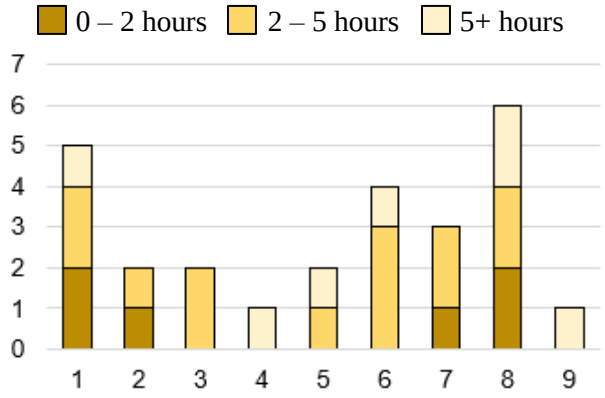
Reason for participating in the oral assessment	
1	To get a better understanding; to reinforce my learning
2	To test my own learning, prove to myself
3	To more accurately represent my knowledge by explaining to an examiner
4	To understand where my knowledge is lacking
5	To get clarification on parts I'm struggling with
6	To have an opportunity to be forced to review
7	To experience an oral assessment
8	To get the extra credit

Reason for not participating in the oral assessment	
9	Because the extra credit is minimal
10	Because I do not feel ready
11	Because I feel there is no benefit in terms of learning versus stress
12	Because I accept the consequences of not doing well on the written exam
13	I'm too busy with other classes
14	Because of external reasons (sickness, other obligations)
15	I accidentally missed my slot
16	I accidentally signed up but was not eligible (did not fail the written exam)
17	I was absent for the written exam

Figure 2. Students' reasons for participating or not participating in the oral assessment when they qualified for it (from the pre-survey and post-assessment survey).

Figure 3 shows for the students who did participate in the oral assessment how they prepared and how much time they spent preparing. The how-they-prepared is the coded version of a free

response question from the post-assessment survey. The time estimate is from a multiple-choice question in that same survey. Note that the course was offered as a flipped classroom. In this context, “rewatching pre-lecture videos” refers to revisiting the theoretical foundations of the course (15% of the students). In addition, 38% of the students listed either visiting office hours (a goal of our intervention) or adopting strategies specifically tailored towards oral assessments (#8 and #9) that guide them toward higher order reasoning skills. While this is encouraging, it also suggests that even more explicit guidance is valuable on how to best leverage the opportunity of an oral assessment and how to approach learning in general. For these conclusions, we want to be mindful that they are based on a limited number of data points.



“How did you prepare for the oral assessment?”	
1	Solving the oral assessment questions.
2	Same as preparing for the written exam.
3	Doing homework problems.
4	Revisiting the written exam.
5	Going through my notes.
6	Rewatching pre-lecture videos
7	Going to office/tutoring hours.
8	Focusing on how to explain my thoughts.
9	Doing a mock exam with others.

Figure 3. How and how much time students spent preparing for the oral assessments (post-assessment survey, N = 26).

To further explore the impact of the intervention on student learning, we included a series of questions related to this in both the post-assessment and end-of-term surveys. These results are summarized in Figures 4 and 5. Responses were on a 5-point Likert scale, encoded -1 through 1 (step size of 0.5), with the average reported for each question in the tables. All answers are for students who participated in the oral assessment. The one exception is Figure 4(d), which was asked of all the students who did not qualify for the oral assessment. This was added to contrast it to question (c) in the same figure. It suggests that taking the oral assessment resulted in students seeing more intrinsic value in it.

In addition, these results also indicate that the oral assessment had a positive effect on students in terms of their awareness of where they struggle, their motivation to learn and their confidence in themselves. These are all important factors contributing to the students’ self-efficacy, which in term has been reported to have a positive impact on learning [24, 25]. Also, as supported by the results on student confidence in Figure 5d, by not framing the intervention as remedial we avoided the negative effect on student self-perception that Vandenbussche et al observed in their study [20].

"Please indicate how much you agree or disagree with the following statement:"		Average
(a)	"Because of the oral assessment, I am more aware of which concepts I am struggling with."	0.68 (N = 26)
(b)	"Interaction with a Prof/TA/Reader during the oral assessment has increased my motivation to learn."	0.58 (N = 26)
(c)	"I would have participated in the oral assessment, even if it did not involve any extra credit."	0.39 (N = 26)
(d)	*"I would have liked to be able to do an oral assessment, even if it did not involve any extra credit." [for students who did not qualify for the oral assessment]*	-0.05 (N = 162)

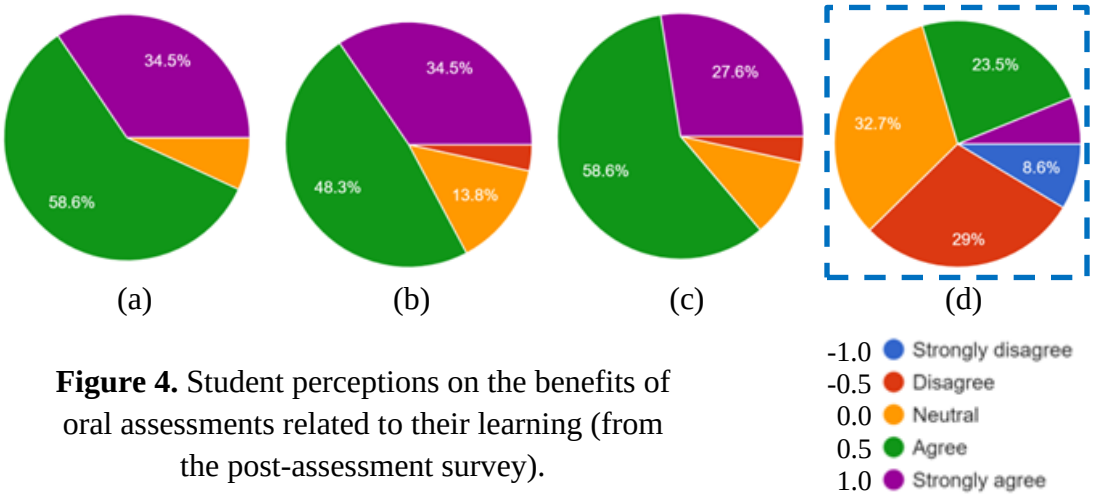


Figure 4. Student perceptions on the benefits of oral assessments related to their learning (from the post-assessment survey).

"Please indicate how much you agree or disagree with the following statement:"		Average
(a)	"Preparing for the oral assessment increased my conceptual understanding of the subject matter."	0.61 (N = 18)
(b)	"Doing the oral assessment increased my conceptual understanding of the subject matter."	0.64 (N = 18)
(c)	"The oral assessment in this course has changed my studying strategy for learning."	0.28 (N = 18)
(d)	"Interactions with the assessor during the oral assessment positively contributed to my confidence in myself and my abilities."	0.61 (N = 18)

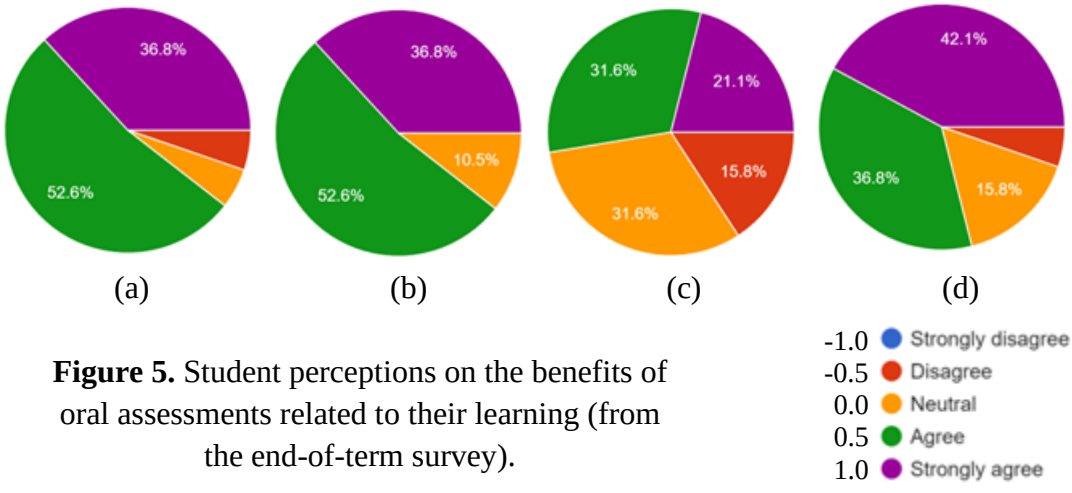


Figure 5. Student perceptions on the benefits of oral assessments related to their learning (from the end-of-term survey).

Finally, Figure 6 compares how students rate the various course elements in terms of how they contributed to their learning. This data was obtained from the end-of-term survey, with averages reported from a 5-point Likert scale multiple choice question. Note that students who participated in the oral assessment rank it as the highest contributor. This further illustrates the value of the oral assessment as a tool to directly aid students in their learning.

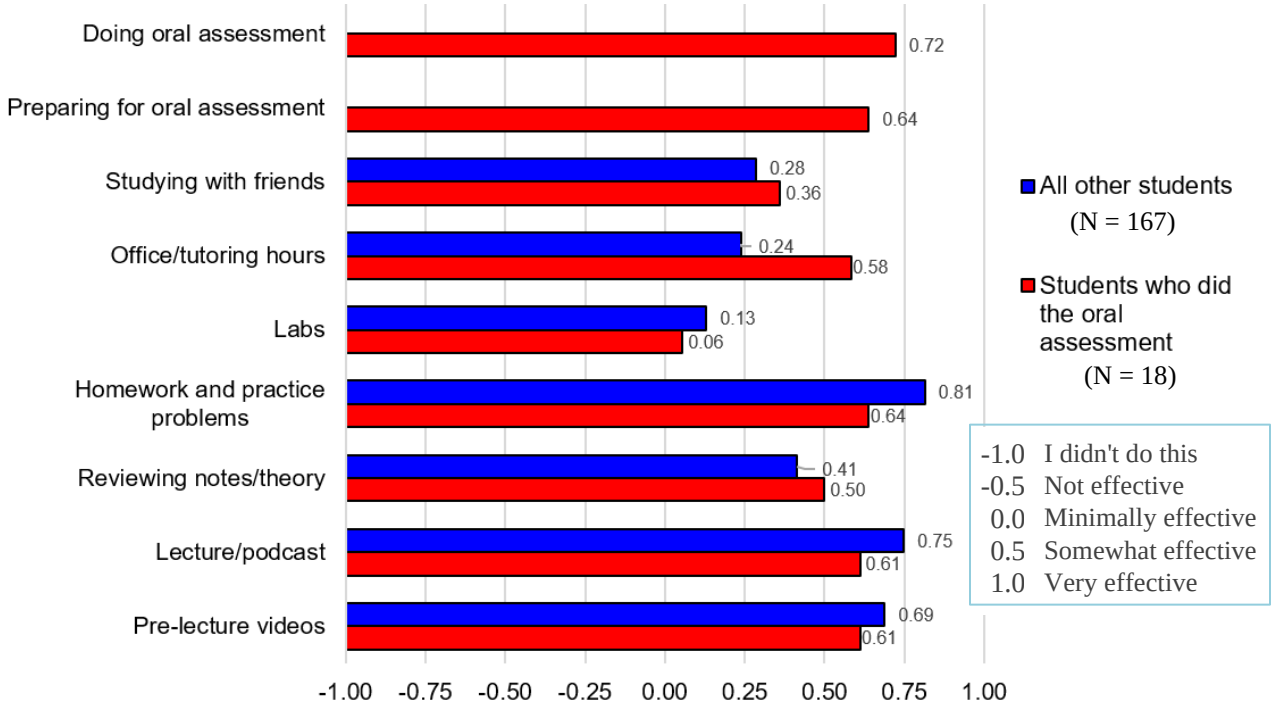
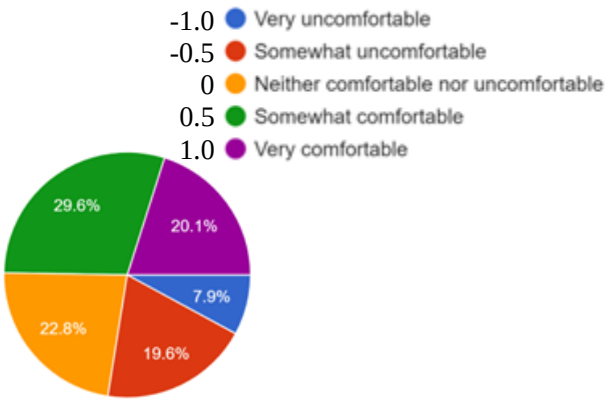


Figure 6. Student responses to the question "Rate how the following class elements contributed positively to your learning." (from the end-of-term survey).

Relatedness

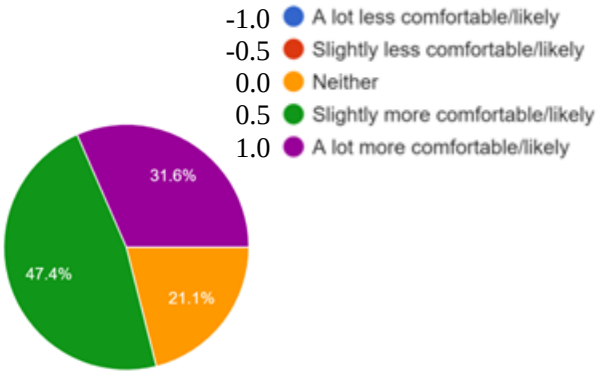
In addition to the impact on learning, we also explore if the oral assessment impacts a student's sense of belonging, and specifically whether students feel more connected to the instructional team and are more willing to visit tutoring/office hours to seek help. To serve as a benchmark, we asked students in the pre-survey about their current comfort level in reaching out for help. The results of this multiple-choice question on a 5-point Likert scale are shown in Figure 7. Note that some students report being less than comfortable. This is even though efforts were made to create a welcoming environment (tutoring/office hours were plentiful and in an open space tutoring center, instructional assistants invited students during the first lecture and the flipped classroom is known to have a positive effect on belongingness [34]).



"How comfortable do you feel about reaching out to the instructional team for help? (such as through emails, asking questions in class or attending office hours)?"	
Group	average
All students (N = 189)	0.17
Students who are planning to take the oral assessment (N = 36)	-0.02
Students who are not planning to take the oral assessment (N = 15)	0.14

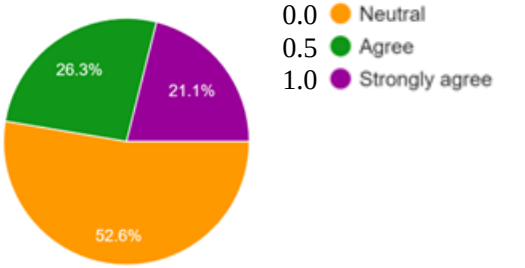
Figure 7. Initial student comfort level reaching out the instructional team (from the pre-survey).

To evaluate the effect of the oral assessment, we asked students two related questions on the end-of-term survey, regarding their comfort level after the assessment as well as their self-reported attendance to tutoring/office hours. These results are shown in Figures 8 and 9. Both show a distinct positive impact.



"Did taking the oral assessment make you more or less comfortable to reach out to the instructional team for help (such as in office hours, via email, or through other methods)?"	
Group	average
Students who took the oral assessment (N = 19)	0.59

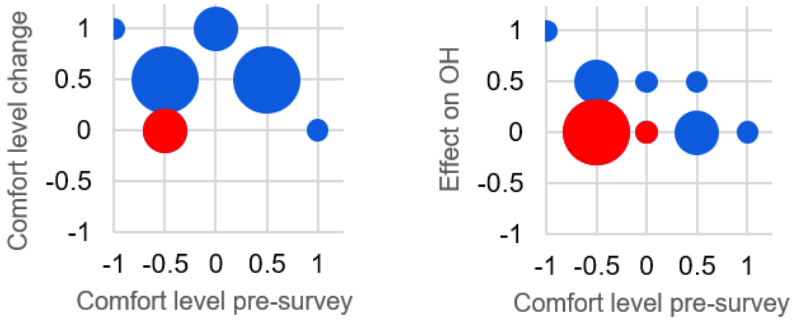
Figure 8. Student comfort level change reaching out the instructional team (from the end-of-term survey).



"Please indicate how much you agree or disagree with the following statement: Because of the oral assessment, I came to tutoring/office hours more often for the remainder of the quarter."	
Group	average
Students who took the oral assessment (N = 19)	0.36

Figure 9. Student self-reported change in participation in office/tutoring hours (from the end-of-term survey).

We anticipated that students who already were comfortable in coming to office hours would report less of an effect of the assessment. To capture this effect, we looked at the correlation between comfort level as reported in the pre-survey (Figure 7) versus their change in comfort level (Figure 8) and self-reported participation in office/tutoring hours (Figure 9). These results are shown in Figure 10, with the area of the bubble corresponding to the number of responses (N = 12; the number of students who took the oral assessment and did the pre-survey and the end-of-quarter survey). These results show that for some students (red bubbles) the intervention had no effect where one would hope there would be one. For all other students (blue bubbles), there was either a positive effect (increase in comfort level/office hour attendance) or no effect when comfort level was already high.



(a) (b)

Figure 10. Correlation of pre-survey comfort level with change in comfort level and effect on office/tutoring hours.

Exam Performance

Finally, we also want to consider whether these positive effects on learning, relatedness and self-efficacy translate to increased performance in the course. To this end, we looked at the student performance on the final exam, which was a 3-hour written exam. The reported data considers students who participated in both the first written exam (which was used to identify at-risk students who would be invited for the oral assessment) and the final exam.

For our study, we approached the oral assessment as an early intervention opportunity. However, as extra credit was offered as an incentive, a grade for the oral assessment was given (scored out of 2). Figure 11 plots the relationship of this grade versus their score on the final exam, showing a loose correlation, which confirms that the oral assessment has some value as an evaluation of and motivator for conceptual knowledge [16]. However, because the oral assessment was integrated as an early intervention pathway, significant value add occurs *after* the oral

assessment, in terms of changes in study approach, self-efficacy, relatedness and increased comfort in reaching out for help.

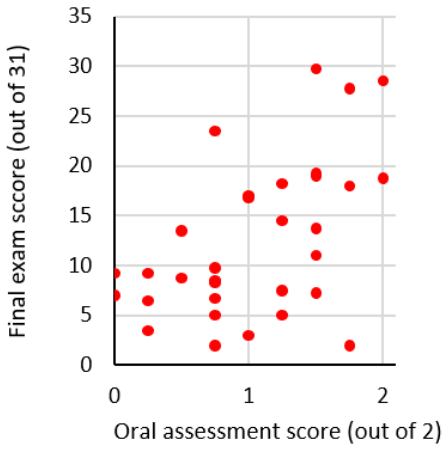


Figure 11. Oral exam score versus final exam score.

To study these effects, Table 2 lists the scores on the first written exam (which was also used to identify at-risk students) and the final exam. The gain metric captures how much better they scored on the final compared to the first written test (i.e., before the intervention). Not unexpectedly, students who passed the first exam had a lower gain. This can be explained by the fact that they have less room to improve and therefore are statistically more likely to score worse. In comparing the students who were eligible for the oral assessment but decided to take it or not take it, we see a higher gain for those who did. However, it is difficult to rule out an impact from self-selection bias. Related work also suffered from this issue [20, 21].

Table 2. Comparison of the scores on the first written exam and final exam.

	Exam	mean	standard deviation
Student who did the oral assessment (N = 32)	First written exam score	27.02%	13.77%
	Final exam score	40.17%	25.10%
	Gain	13.15%	20.31%
Students who elected not to do the oral assessment (N = 28)	First written exam score	25.45%	15.77%
	Final exam score	34.77%	22.05%
	Gain	9.32%	20.75%
Students who were not eligible for the oral assessment (N = 268)	First written exam score	71.86%	29.19%
	Final exam score	63.72%	25.79%
	Gain	-8.14%	22.81%

To get another data point regarding potential benefits on student performance, we also considered a comparison group. Specifically, we looked at the same course, offered by the same instructor, in a prior term (Winter'22). The course structure was the same, except that the proposed intervention was not used. A total of 173 students were enrolled during that term. To account for unequal exam difficulty levels between Winter'22 and the term of our intervention (Fall'22), we propose to use normalized student scores for all exams:

$$x_{norm} = \frac{x - \mu}{\sigma}$$

In this equation, x is the original exam score of a student, μ the average exam score in the course that term, and σ the standard deviation. Essentially, a student's normalized score is the number of standard deviations they are above the class average for that exam. As a

comparison group, we consider the students who would have been eligible for an oral assessment in Winter'22, had the intervention been implemented then. The cutoff in terms of normalized first written exam score in Fall'22 was -0.75, and we used this same cutoff to create the comparison group for Winter'22. This means that for both terms, the group of "oral assessment eligible students" was those who were 0.75 sigma or more below the average. This represented a similar fraction of students. Table 3 provides the normalized exam performance and gain from the first written exam to the final, for both groups. We notice that the normalized gain is higher for the students who took the oral assessment, compared to those who did not as well as the students in the Winter'22 comparison group (0.62 versus 0.47 and 0.43 respectively). Unfortunately, the sample sizes are too small to reach statistical significance.

Table 3. Normalized scores versus those of a comparison group from a prior term.

	Exam	mean	standard deviation
Student who did the oral assessment (N = 32)	First written exam score (normalized)	-1.54	0.47
	Final exam score (normalized)	-0.91	0.97
	Gain (normalized)	0.62	0.79
Students who elected not to do the oral assessment (N = 28)	First written exam score (normalized)	-1.59	0.54
	Final exam score (normalized)	-1.12	0.85
	Gain (normalized)	0.47	0.79
Students who would have been eligible for the oral assessment in Winter'22 (N = 43)	First written exam score (normalized)	-1.34	0.34
	Final exam score (normalized)	-0.91	0.89
	Gain (normalized)	0.43	0.86

Finally, we also investigated if increased relatedness influenced exam performance. Figure 12 shows the increased level of comfort with members of the instructional team (as measured by the end-of-term survey, see Figure 8) versus exam score improvement (gain from Table 3), for students who participated in the oral assessment in Fall'22. The colors encode the self-reported increased attendance to tutoring/office hours (as measured by the end-of-term survey, see Figure 9). The general trend seems to suggest that students who felt more comfortable and went to office hours more, saw a higher impact on their performance. However, as we have limited data points (N = 14; students who did the relevant survey and both exams), it is difficult to draw any definitive conclusions.

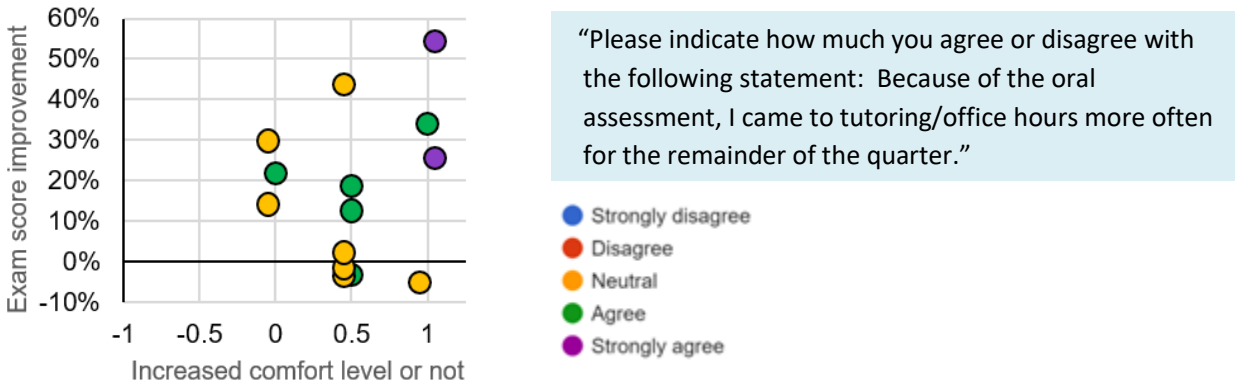


Figure 12. Exam performance versus increased relatedness.

Conclusions

The goal of this study was to explore whether oral assessments could be used as an effective early intervention pathway, as they offer both affective and learning benefits. Specifically, we explored how they improve student motivation, self-efficacy, and relatedness to the instructional team, which ties to office/tutoring hour attendance Leveraging oral assessments in an early intervention strategy for at-risk students, focuses them where they are most impactful and automatically addresses their main implementation challenge, i.e., scalability.

We implemented this idea in a lower division electronics course. Despite the overall class size of around 300 students, the intervention required only 49 slots of 20 minutes (fewer than 17 hours, distributed between members of the instructional team). While still a significant time investment, we believe it is worthwhile as it specifically targets students who can benefit the most. In addition, as fewer students participate in the oral assessments, one could consider allowing more time per student (currently, it was limited to 20 minutes), to increase the impact as an early intervention strategy, or to offer the opportunity after every written test.

The data gathered is encouraging, suggesting gains in self-efficacy and connectedness, and potentially performance. However, a side effect of creating an approach that improves the scalability of oral assessments, is a smaller sample size for our study. As such, our results, while suggesting positive outcomes, did not reach the level of statistical significance. A similar issue was encountered by related work on early intervention strategies [19-21]. Repeating the intervention in additional/future courses is necessary to verify if these initial observations can be confirmed. Also, as our sample sizes were small, we did not further subdivide our results by assessor (instructor versus the instructional assistants) or according to gender, ethnicity, or other demographic factors, as these subdivisions would lack any statistical significance.

Finally, student impressions of the oral assessment were very positive. On the end-of-term survey, we included this concluding optional question for students who participated: “What do you feel were the benefits of the oral assessment if any? What worked well for you?”. The responses we received are below:

- “I learned concepts I was not sure of such as passive sign convention.”
- “clarity on knowledge”
- “The tutor gave me a list of topics that I needed to review and helped me feel more comfortable about going to tutoring”
- “The oral assessment pushed me to reach out to staff for assistance and further help in making sure I understood the material.”
- “I was able to get things cleared out and felt comfortable when the assessor treated my mistakes and confusions as normal considering I was not very confident. I also felt like a bit more relieved knowing I was also improving my grade even if it was a little bit and in return I am also improving my understanding of the material.”
- “Knowing i was given a second chance gave me the push to study the concepts harder than before, as otherwise i would have felt i was behind and demoralized by myself. It also gave me comfort in reaching out to the tutors going forward.”

Even though these are anecdotal, it is encouraging to see intervention goals of self-efficacy and relatedness echoed by the students.

Acknowledgements

This material is based upon work supported by the National Science Foundation under Grant No. 2044472. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation. We would like to thank our colleagues Dr. Celeste Pilegard, Dr. Nathan Delson, Dr., Dr. Carolyn Sandoval, Josephine Relaford-Doyle, and Dr. Mia Minnes for the helpful discussions, and our undergraduate and graduate instructional assistants for helping administer the oral assessment.

References

[1] M. L., Simpson, C. R. Hynd, S. L. Nist, & Burrell, K. I. (1997). College academic assistance programs and practices. Educational Psychology Review, 9(1), 39-87.

[2] K. Durkin and A. Main, (2002). Discipline-based study skills support for first-year undergraduate students. Active Learning in Higher Education, 3(1), 24-39.

[3] R. E. Mayer and M.C. Wittrock (2006). Problem solving. Handbook of educational psychology, 2, 287-303.

[4] G. Joughin, A short guide to oral assessment. Leeds Met Press in association with University of Wollongong, 2010.
[5] P. Sotiriadou, D. Logan, A. Daly, and R. Guest. "The role of authentic assessment to preserve academic integrity and promote skill development and employability." *Studies in Higher Education* 45, no. 11 (2020): 2132-2148.
[6] B. B. Jensen, "Oral assessment in engineering education." *International Journal of Electrical Engineering Education* 47, no. 4 (2010): 375-379.
[7] K.P. Rouser, "Oral assessments of student learning in undergraduate aerospace propulsion and power courses." *Journal of Engineering for Gas Turbines and Power* 139, no. 12 (2017).
[8] M. Huxham, F. Campbell, and J. Westwood. "Oral versus written assessments: A test of student performance and attitudes." *Assessment & Evaluation in Higher Education* 37, no. 1 (2012): 125-136.
[9] P. Iannone, and A. Simpson. "Students' views of oral performance assessment in mathematics: straddling the 'assessment of and 'assessment for learning divide." *Assessment & Evaluation in Higher Education* 40, no. 7 (2015): 971-987.
[10] A. P. Dicks et al. "Undergraduate oral examinations in a university organic chemistry curriculum." Journal of Chemical Education 89.12 (2012): 1506-1510.
[11] J. Douglas and R. Knighten. "Using Oral Quizzes in an Engineering Mechanics Course." 2014 ASEE North Midwest Section Conference. (2014)
[12] Y. Zhao, "Impact of Oral Exams on a Thermodynamics Course Performance." Proceeding of 2018 ASEE Zone IV Conference, Boulder, Colorado, 2018
[13] D. Grunwald, E. Boese, R. Hoenigman, Andy Sayler, and Judith Stafford. "Personalized Attention@ Scale: Talk Isn't Cheap, But It's Effective." *Proceedings of the 46th ACM Technical Symposium on Computer Science Education*, pp. 610-615. 2015.
[14] C. Schurgers, S. Baghdadchi, M. Lubarda, M. Ghazinejad, A. Phan and H. Qi, "Evaluating Oral Exams in Large Undergraduate Engineering Courses," 2021 IEEE International Symposium on Antennas and Propagation and USNC-URSI Radio Science Meeting (APS/URSI), 2021.
[15] L. Roecker, "Using oral examination as a technique to assess student understanding and teaching effectiveness." Journal of Chemical Education 84, no. 10 (2007): 1663.
[16] M. Kim, C. Pilegard, M. Lubarda, C. Schurgers, S. Baghdadchi, A. Phan, H. Qi, "Midterm oral exams add value as a predictor of final written exam performance in engineering classes: A multiple regression analysis", ASEE Annual Conference & Exposition, Minneapolis, Minnesota, 2022.
[17] J. Clark et al. "Peer Mentoring Program: Providing early intervention and support to improve retention and success of women in engineering, computer science, and physics." 2015 ASEE Annual Conference & Exposition. 2015.
[18] R. H. Fenske et al. "Early Intervention Programs. Opening the Door to Higher Education," ASHE-ERIC Higher Education Report Vol. 25, No. 6. ERIC Clearinghouse on Higher

Education, The George Washington University, One Dupont Circle, Suite 630, Washington, DC 20036-1183, 1997.
[19] J. Lavelle and R. Keltie. "Calculus Intervention For First Semester Engineering Students." 2005 Annual Conference. 2005.
[20] J. Vandenbussche, R. Lake and C. Scherrer. "An incentivized early remediation program in Calculus I." International journal of mathematical education in science and technology 49.8 (2018): 1235-1249.
[21] R. M. Bennett, W. Schleter, and D. Raj Raman. "A success enhancement program after the first test in freshman engineering." 2012 ASEE Annual Conference & Exposition. 2012.
[22] L. Bottomley and J. S. Carter. "Assessment of the Effects of Participation in a Summer Bridge Experience for Women." 2020 ASEE Virtual Annual Conference Content Access. 2020.
[23] H. Hazen "Use of oral examinations to assess student learning in the social sciences." Journal of geography in higher education 44.4 (2020): 592-607.
[24] B. J. Zimmerman, "Self-efficacy: An essential motive to learn." Contemporary educational psychology 25.1 (2000): 82-91.
[25] G. Hackett and N. E. Betz. "An exploration of the mathematics self-efficacy/mathematics performance correspondence." Journal for research in Mathematics Education 20.3 (1989): 261-273.
[26] L. Ritter et al. "A Study of Student Perceptions of Office Hours." Journal on Excellence in College Teaching 32.4: 81-115.
[27] W. Griffin et al. "Starting the conversation: An exploratory study of factors that influence student office hour use." College Teaching 62.3 (2014): 94-99.
[28] Y. K. Kim and L. J. Sax. "The impact of college students' interactions with faculty: A review of general and conditional effects." Higher education: Handbook of theory and research: Published under the sponsorship of the Association for Institutional Research (AIR) and the Association for the Study of Higher Education (ASHE) (2017): 85-139.
[29] G. D. Kuh, and S. Hu. "The effects of student-faculty interaction in the 1990s." The review of higher education 24.3 (2001): 309-332.
[30] S. L. Dika, "Relations with faculty as social capital for college students: Evidence from Puerto Rico." Journal of College Student Development 53.4 (2012): 596-610.
[31] B.-S. Eylon, and M. C. Linn. "Learning and instruction: An examination of four research perspectives in science education." Review of educational research 58.3 (1988): 251-301.
[32] E. L. Deci and R. M. Ryan. "Self-determination theory." (2012).
[33] M. Parnes et al. "Closing the college achievement gap: Impacts and processes of a help-seeking intervention." Journal of Applied Developmental Psychology 67 (2020): 101121.
[34] M. Jdaitawi, "The effect of flipped classroom strategy on students learning outcomes." International Journal of Instruction 12.3 (2019): 665-680.