

Principles of Programming Languages @ Scale:

The Value of Student Collaboration

Spencer Wilson

University of Colorado Boulder

Abstract. With the growing promise of high income for advanced education in IT fields, many higher academic institutions continue to observe larger admissions in their computer science programs and related fields [1] [2] [3]. Over time, this has led to higher enrollment in courses, limiting the units of one-on-one support available to students, prompting a change in methods to teach effectively at larger scales. Accordingly, many courses have gone to focus on teaching at scale with less regard for the learning value to each student, or continuing their previous teaching methods with limited exploration in adaptation to scale. For each choice, we observe an unwanted trend in unethical behavior from students as we degrade the resources and value posed to students paying for in-person university courses compared against the value they might derive from educational alternatives such as asynchronous or semi-synchronous learning environments. For this paper, we propose that students do not commit acts of academic dishonesty because they are malicious beings. Students commit these acts because they feel that they do not have adequate resources to perform some task through honest means. Here, we propose the use of peer-interviews to scale effective teaching in the Principles of Programming Languages course based on lessons learned from review of literature on effective education at scale.

Keywords— Education at Scale, Computer Science, Principles of Programming Languages, Peer Grading, Ungrading, Interview Grading

1 Introduction

As is standard in academia, there is a lag between increased enrollment and funds for the hiring of additional staff. Over time, we have seen the ratio of staff resources per student decreases. With this resource decreasing, a key value proposition disappears and an important question arises: How do we provide an effective learning experience to our students at scale? In this paper we explore the use of peer-interviewing in an ungraded and self-reflective model for formative assessment on course assignments for CSCI 3155 The Principles of Programming Languages at the University of Colorado Boulder.

We define a method of peer-to-peer reflective interviewing to engage students in a highly scale-able manner that improves student agency in learning. We conducted interviews with students that completed CSCI 3155 in Fall 2022 in early 2024 and use these interviews to suggest personas that take the course at the University of Colorado Boulder. We use these personas to suggest how students will be impacted by our change to the interview process by four dimensions:

1. What impact does this have on personas' interactions with other students?
2. What impact does this have on personas' ability to correctly assess their own performance?
3. What impact does this have on the personas' performance?
4. What impact does this have on personas' ethical behavior in the course?

2 Background

2.1 Cost

As the scale of classrooms increase, some suggest that the value for students is decreasing. So what is the cost to take this course for a student? What exactly are they paying for and why should pay for do this as opposed to learning through an online degree program or a Coursera course? In review of the university data we see that the average fee collected by students in the course is \$3,369.25. So what support does

a student get for their three-thousand dollars paid to the university? Supposing the student meets the university standard of 3 hours of work per credit hour, a 4 credit hour class like CSCI 3155 will be 12 hours per week across 16 weeks. Per hour of student engaging with course content, the individual pays on average \$17.55. They are paying for the full package at a 'reasonable' cost. It's not only about coming to lecture, but it is also about having a constructive environment that challenges them to grow their mastery of the material through reasonable assignments and additional resources that support this growth. Suppose that homework content is freely available online like it was for CSCI 3155 in Fall 2022, then what are they paying for? They are paying for the social environment. They are paying for the emotional and technical support provided to them by the course staff and other organizations in the university as they choose whether to participate in lectures, recitations, office hours, and online forums such as piazza. Of course there are asymmetries in how students use the direct support provided to them with a small handful of students using the majority of resources (active in lecture, recitation, piazza, office hours, and other factors in their social network related to the course) while a large group of students do not leverage these resources. Accordingly, it is important for the staff of a course to help students understand the cost of resources that they are paying for and encourage all students to leverage what is available to support their learning.

2.2 Effectiveness

For this paper, effectiveness in learning refers to providing experiences that engage the student and enables high levels of cognition as defined by the Bloom's taxonomy.

Bloom's Taxonomy In this paper we center on the Bloom's taxonomy of learning as the measure of student achievement in mastering the course material. The popular 2001 revision to Bloom's taxonomy suggests a linear progression of cognition from "remember", "understand", "apply", "analyze", "evaluate", and "create" [4]. Here, "remember" is the lowest level of cognition that a student can achieve, in which they know a few seemingly disparate facts. On the other end "create" is the

highest level of cognition, in which students can build on all they have learned to form well-reasoned solutions to complex problems which are novel to the learner. While "create" is rarely the goal of an assignment and it is not a reasonable goal for all students enrolled, it is an ideal goal for the course as a whole for each student that are adequately resourced to complete the course (students that have the time, energy, and background knowledge to succeed).

Ungrading Next, we explore ungrading models, by which we move away from a model of grading out of one hundred points and toward a model of "X", "✓-", "✓", "✓+", or some other naming model to represent a distinction from work that is unacceptable (X,✓-) versus "good enough" (✓), or even exceptional (✓+). In various un-grading models such as reflective un-grading, contract grading, and standards based grading we move the staff focus away from time obsessing over the difference in grade from an 85% to a 88%, and instead state, that's a "✓". This allows us to re-prioritize substantive feedback for our students [5] [6] [7]. While this requires constant buy-in from the course staff and students to ensure success across the term as students become co-conspirators in this different educational model, the model has proven effective in many college courses including upper division topics [7] [8] [9] [10]. This concept can be leveraged effectively in interview grading to emphasize formative feedback over a course grade for the student, helping to move students toward intrinsic learning rewards over extrinsic ones [11].

Interview Grading A tool for effective instruction explored at various institutions is to give students an oral assessment of their work called "interview grading". In interview grading, students **evaluate** their mastery of the course material with an oral review of their written assignments. Interview grading has been shown to hold value for students being accountable to their own learning. It works best in a small class setting where the instructors can manage all of the interviews [12] [13]. However, it can be done at scale, by offloading the effort to support staff such as graduate teaching assistants and graders [14]. It is important to note that as proposed, this doesn't continue to scale well as more students means more time for grading by "expert" course staff.

Reflective Learning In reflective learning, we ask students to have agency in their own education and continuously reflect on what they have learned, what they are struggling with, and how they could potentially apply what they have learned to reach their own goals. In fact, there is a model of un-grading built around this concept, sometimes called "reflective un-grading" or "big-U Un-gadding" [5]. Here we develop a learning environment where students must author self-reflections and even recommend their own grade for the course. We as course staff might then decide if the students' self-reflection and decided grade is accurate, or how it differs and discuss significant differences with the students. Alternatively, to increase the scale-ability of this model, the course staff can trust the validity of the student assessment and instead analyze the student reports to understand what students are doing well in then leverage that knowledge to improve future lectures and readings based on the student experience.

2.3 Scale-ability

For this paper, scale-ability in education refers to providing consistent learning opportunities to as many students as possible. Some obvious places to look for scale-able education tools are the use of artificial intelligence (Ai) in the classroom, and the world of online learning [15] [16] [17]. While Ai in the classroom is promising, it is currently burdensome to implement, so we'll focus more on tools from online learning. What is found to be most important in scaling education online is encouraging collaboration between students in peer-to-peer interactions. After all, more students in the classroom means more students that can interact with other students. Beyond technology integration's, this is the most scale-able resource for the course as enrollment increases. Let us explore two key tools in improving peer-to-peer interactions.

Peer Grading Having students grade each other is considered a must for effective online education at scale. While many students are resistant to peer grading and do not believe it to be as helpful as feedback from their course staff, it has been shown to be effective[15]. This scales infinitely, as more students yields more people to perform the reviews. Perhaps the most important aspect of doing this effectively at scale is to

have a way of assessing the students review capabilities. The literature suggests an effective method to ensure effective peer grading is to have some kind of training assignment. Here, students complete an assignment to demonstrate acceptable knowledge of the peer review process early in the semester [18]. This method has been employed extensively in the online learning environment where scale is nearly limitless.

Discussion Forums Additionally, to increase a sense of belonging and community in a large class - be it online or in person - we see a recommendation for online discussion forums such as Slack, Discord, Piazza, and Zulip [15] [19]. Here many students are able to engage with the material and start discussions with their peers. It is best practice to have course staff monitor and collaborate on this forum as well. While this requires some time from staff to manage the forum, this is often worth the effort for larger sized classes as it engages students on some semi-synchronous forum where they can ask questions and discuss topics beyond the confines of class time.

2.4 The current syllabus

The current course syllabus has seen continued decrease in effectiveness over the years as course enrollment increases - anecdotally. The following assessments are used to construct a course that in practice is shown to be highly effective with seventy students. However, it is struggling to stay effective at one-hundred-fifty students and does not look promising for three-hundred students:

1. Lecture Preparations: be it assigned reading, attempting portions of the lab, or watching video lectures on key topics, lecture preparation is the key means for students to **remember**, **understand**, and begin to **apply** their knowledge of the related material.
2. Participation: a formative assessment in which students **analyze** information through discussions during in-person class sessions, building on lecture preparations.
3. Labs: a formative assessment in which students **analyze** topics of interest which serves as the basis of student learning. All students complete the same lab in teams of two to three students and use their findings in the assignment to engage class

discussions on the related topics. The lab is auto-graded for correctness against a set of pre-defined tests which are partially shared with the students.

4. Grading Interviews: a formative assessment in which students **evaluate** their mastery on the lab material with twelve minute staff-interviews one-on-one with a member of the course staff in an ungraded X/√+ style score returned with limited personalized feedback and a score out of one-hundred percent. This interview is graded on the basis of student's ability to correctly answer the questions in the interview within the time provided.
5. Exams: The course is composed of one midterm exam and one final exam. These summative assessments require students to **analyze**, **evaluate**, and **create** novel solutions to relevant problems in a timed assessment that is manually graded by the course staff and returned to students with limited qualitative feedback at the midterm and no feedback at the final.

Grading Interview with course staff The following describes the structure of interview grading for CSCI 3155 in Fall 2022. We will later discuss the reality of a student experience in these staff-interviews. Each member of the course staff conducting interviews has an assured "**analyze**" level of learning on the material and is variable in their level of mastery and their interest in hosting an effective learning environment at each interview [14]. At the end of each lab, each student sings up for a single staff-interview with a member of the course staff. The student attends the interview without prior knowledge of the questions that will be asked, and perform the interview in a twelve minute slot. At the end of the interview, the course staff, tells the student how they performed on each question in an "X, √-, √, √+" scale and works with the student in the time available to discuss plans for improvement as necessary. The course staff also takes time to celebrate what the students have already mastered and encourage their continued success. The member of the course staff are able to pivot the interview as needed to ask follow-up questions of the student in the Socratic method that encourages the student to create a more comprehensive understanding of the related topics.

3 Experiment

3.1 Interview Design

In this experiment we conducted interviews with Fall 2022 students of CSCI 3155 at the University of Colorado Boulder. We designed consent forms detailing the process for the research study and the data that will be collected informed by Citi training requirements to minimize potential harm, maximize potential benefit, collect informed consent, and detail the process to collect and distribute data collected through the research study. Do to an inability to reach all past student, we sent emails to ten students that we personally recall working with, that represented a wide range of mastery with the course concepts. From the ten candidates solicited, seven did not respond, and the remaining three agreed to participate in the study completing two interviews each on their experience in CSCI 3155 Fall 2022 and topics related to CSCI 3155, interview grading, and peer interactions. The interview questions reviewed are as follows:

1. As a student in 3155, what was your experience working with a "lab partner" on your labs? Please elaborate.
2. As a student in 3155, you had interview grading with a member of the course staff. What was your typical process to prepare for this interview? Did it evolve over the course of the semester?
3. As a student in 3155, what was your experience working with a "grader" for your interviews at the end of the lab? Please elaborate.
4. What impact, if any, did the interviews have on your confidence to succeed in the course?
5. Do you have experiences with interview grading in other courses? If so, how did those experiences compare to your experience in 3155?
6. What would you describe as your learning goal for 3155?
7. What would you describe as your ability to achieve that learning goal?
8. What impact, if any, did interview grading have on your relation to that learning goal?

9. Tell me about your favorite course in recent memory. What aspects of that course were positive for you? (We won't record the course name/title as it could potentially be used to identify you.)
10. Tell me about your most recent positive experience in a course that had 200 or more students. In particular, a course where the full course was positive and not only a singular event in the course. What made it a positive experience? (We won't record the course name/title as it could potentially be used to identify you.)
11. Tell me about your most recent negative experience in a course that had 200 or more students. In particular, a course where the full course was negative and not only a singular event in the course. What made it a negative experience? (We won't record the course name/title as it could potentially be used to identify you.)
12. This research study is exploring the challenges faced in CSCI 3155 as we continue to see larger enrollments over time. We are exploring ways to provide value to our students at these larger scales. As a student in Fall 2022 CSCI 3155, what aspects of the course helped in your ability to meet your goals? (FOLLOWUP: What aspects of the course impeded your ability to meet your goals?)

DISCLAIMER Some interview questions had slight changes between interview events in an attempt to removed any bias that may have been present in the questions. Specifically questions 1 and 3 were phrased as follows for the interview with participant 13 before being changed for all future interviews:

1. Q1: As a student in 3155, what was your experience working with a "lab partner" on your labs? **Was it positive? negative? mix?** please elaborate?
2. Q3: As a student in 3155, what was your experience working with a "grader" for your interviews at the end of the lab? **Was it positive? negative? mix?** please elaborate?

DISCLAIMER The researcher had personal connections and discussions with participants between when the participant took CSCI 3155 in Fall 2022 and when the interviews were conducted. Accordingly, it is possible that the research participants presented information during their interviews that they believe the researcher wanted

to hear rather than stating their honest opinions. When language was used to that effect, then the researcher would explain that no such attempts are necessary as we are looking to gather objective data about the the participant’s personal experience.

3.2 Proposal

In this experiment we propose one core change to the course syllabus. Here, peer-interviews are not graded based on the students’ correct answers to the interview questions, but instead purely on the students completion of a reflection of their peer-interview process and their current understanding of the material. We emphasize the formative nature of the interview and focus on giving students qualitative feedback on their performance in an ungraded model. We propose a single method of interviews in which students perform the interviews in a peer-to-peer model without staff review. We then require that students complete self-reflections on the peer-interview experience for review by course staff with reactionary interventions as necessary.

The course is comprised of six labs which build off the knowledge of the previous lab. In each lab we propose that students complete a peer-interview and submit a reflection on that process for review by the course staff prior to the beginning of the next lab.

3.3 Interview Process

The new peer-interview process contains four phases

1. Training Phase
2. Interview Phase
3. Reflection Phase
4. Action Phase

Training Phase The training phase is required at the beginning of the semester and is reassigned as needed to students throughout the semester to re-commit the student to this interview grading process. In the training phase, students are given a

detailed introduction to the "X, ✓-, ✓, ✓+" grading method and an explanation of the reasoning for using this method of grading. Students are then given a series of videos on mock-interviews with a grading rubric for the interview using an "X, ✓-, ✓, ✓+" grading system for the topics in the interview. Students are asked to grade the interviewee against the rubric and submit their solutions to an automated grading tool which compares the students' proposed grades to the known grade of the mock interview. While this effort would have large upfront cost, this effort has been shown in other studies to provide great value in reducing overhead throughout the semester by setting clear expectations for students early in the semester [18].

Interview Phase For each lab, the course staff prepares an interview document containing questions, alternate questions, hints, solutions, and grading guidelines. Consider hypothetical students Ethan and Ayden have just completed the lab as student team. Ethan and Ayden each play the role of interviewer and interviewee and should take turns asking each question from the interview document. If the interviewee cannot answer the question sufficiently, then interviewer should review the hints for the question and attempt to give scaffolding that supports the interviewee to a correct solution to the question. If the neither member of the team can understand the question, then the student team should study before continuing the interview. In studying, Ethan and Ayden might collaborate or study individually based on their unique learning styles. This might include resources such as conversational Ai, reviewing the provided solutions to the question, and contacting peers or course staff through resources such as the online discussion forum (piazza) or back-channel student forum (discord). When recommencing the interview, the student team should consider the alternate question rather than the original question, the solutions for which have likely already been reviewed by the student team. In completing the interview, the students should agree on grading for each participant based on their total understanding of the course content exposed through the interview and not only the students initial answer to the question. Here, we encourage student collaboration in learning and increase student autonomy in resourcing their learning when compared to a 12 minute staff-interview that are

closed resources beyond the individual student and the staff member conducting the
interview. Additionally, as the staff is no longer conducting the staff-interviews, they
have more time to dedicate toward supporting student learning in other ways.

Reflection Phase

Student Reflection and Action Planning Ethan and Ayden now meet to review
their performance on the peer-interview and the lab content as a whole. Students
are encouraged to spend about thirty minutes on this exercise. They identify their
performance on a selection of key skills used in the lab and develop a personal action
plan for what they might focus their efforts on in the next lab, taking advantage of
the benefits of reflective learning. While the action plan is personal to the individual,
students are encouraged to collaborate in generating ideas for their action plan. Each
student submits this via a survey form that allows for the aggregation of student data.

Staff Reflection and Action Plan While each member of the course staff has access to
the student reflections, one delegate of the course staff analyzes the student performance
as described by the students in their self-reflections. This data analysis is targeted
at understanding current student strengths and opportunities for improvement. The
delegate presents their finding to the rest of course the staff and the team collaborates
on a plan of action that will build on those stores of knowledge to scaffold learning on
the topics where students have the most opportunity to grow in their learning journey.
In practice, we expect this will take a few hours by the delegate staff member for each
lab as well as a two-hour meeting with the full course staff after the completion of the
interview phase for the lab. For best results, we recommend that this work is completed
as early as possible during the next lab so that the feedback can be leveraged in a timely
fashion.

Action Phase In the action phase, the course staff executes on their plan for im-
proving the course lectures based on common findings in students' gaps in knowledge.
In an attempt to increase transparency of the process and build our students as con-
spirators to the method, the course instruction includes anonymous quotes from the

student reflections and openly recognizes why we are covering certain topics in more depth. The students are also encouraged to act on their own action plans and seek whatever assistant or materials they may need. Toward enabling the students' success, the course staff is listening to students and taking note of what roadblocks exist for the students and actively working at removing those roadblocks wherever staff intervention is necessary while being careful not to remove those critical speed-bumps that students need in order to have autonomy in their own learning and feel the sense of accomplishment.

4 Results

Below we detail the findings from interviews of past students to construct personas that would exist in our proposed framework. We then demonstrate how these personas might be impacted by the change to course structure proposed.

4.1 Interview Findings

In review of the interviews with research participants, we found student responses varied widely, with one participant mimicking the results found by "Personalized Attention @ Scale" that the grader experience was variable, but the interviews were valuable to the student experience both by validating the students learning as well as feeling like part of a community rather than "just a number" [14]. While the remaining participants interviews quickly focused on challenges faced by the students in being able to understand the material from the course and apply it in a way that built their confidence and interest in learning the requisite material.

Through these interviews we found three personas. Each persona is a theoretical description of a student that took the course and not an actual person. The personas described are not yet representative of each student type that exists in an actual course setting. Students individually may represent any of these personas at any point in time throughout the semester. These personas are developed from the data collected in interviews from three different students. Ethan has an "evaluation" level of cognitive

mastery of most content in the course, Ayden an **"apply"** level, and Umbrielle an **"understand"** level. They each approach the course differently as detailed below.

Ethan takes the approach to this course as any other course, to learn the most that he can within the time and other resource constraints that he has each semester. Ethan will occasionally reflect on his goals and abilities without prompting from the course staff and states "[I think that I learn well, but I know that I don't learn everything.]"

Ayden on the other hand, has a different approach. Ayden began the semester with a lot of hopes for the course "to learn how to compile Scala so it will behave as though it were JavaScript" and to learn some "useful things" from the course. At the begging of the semester they had great ambitions for the course but their ability to achieve those goals continued to lag as the semester progressed. They quickly found themselves crowd-sourcing interview questions from their peers, sometimes students like Ethan attempt to provide answers, but often it's other students with similar or lower levels of cognition on the topic that are telling them what questions were asked at the interviews, so "Solicitation was never truly helpful." Ayden made a point to try to learn the material throughout the semester and get help were they could. They often found it difficult to get help from the course staff due to the sheer number of students that wanted help in office hours. They instead found that they received the most help from collaboration with their peers. They remark that where other courses will only imply that students should work together, CSCI 3155 is perhaps the only 150+ student course where the students are explicitly encouraged to collaborate in learning the material for the course. This explicit encouragement created a culture that works well with Ayden's intrinsic motivations to work with other people and build social connections.

Finally, Umbrielle's approach has many similarities to Ayden's approach. However, Umbrielle's learning goal is the same as it is for any other course, she describes it as "my goal was to get an 'A' in the course." Accordingly, she felt that she must do anything she can to pass the provided tests on the lab including finding and using existing solutions from past students. She would also solicit information from her peers as they complete the lab interview to understand what questions will be asked, and

then crowd-source solutions for those questions from her peers. As she explains it, she wanted to learn the material at first, but quickly adapted her approach "[It's not because I wanted to cheat but rather because I did not understand the content well enough.]" She would go on to deepen this pattern throughout the semester to focus more on having correct solutions over understanding the material. By the end of the course she states: "I think 3155 is the only course where I think by the end of the course, I just wanted to pass the class because I was in danger of failing (I think)."

4.2 Example: Lab 2 staff-interviews

Let us detail how these personas would have completed a staff-interview in Fall 2022. In lab 2 students completed an assignment to author an interpreter for a subsection of JavaScript, and interview with a member of the course staff. In the staff-interviews all students are asked about which operators were "overloaded". Among the correct answers, the "+" operator is quickly identified by Ethan as correct. Ethan shares this information with his peers when asked about the interview. When asked about the expression "'hello' + 2 * 5" and its evaluation, Ayden who took time to study for the interview for lab 2 is not able to correctly answer the question. Umbrielle on the other hand, answers the question accurately without understanding why the answer is correct as she also had additional information about the lab interview from one of Ethan's friends.

Here, we demonstrate one challenge of the previous staff-interviews. The staff-interview incentivizes academic dishonesty over learning the lab content. Under the peer-interview model, if done correctly, we expect that lab partners like Ayden and Umbrielle would be able to complete the interview and learn what a correct solution to an **analysis** level question would look like while also having access to information that explains how that solution is derived. Since these students would still not be expected to answer the questions without studying, we may only expect an **apply** level of understanding for the interview.

This decision to have students complete peer-interviews gives trust to the students and poses one significant risk. Now, students that cheat by lying about their completion

of the interview in their interview reflection form will be potentially even worse off then they had been in staff-interviews where cheating often required an **understand** level of cognition for success as they phrase solutions in their own words. Here cheating can be achieved through a **remember** level of knowledge of the material as they can copy and paste words to their reflection without actually completing the peer-interview.

4.3 Peer Interactions

1. What impact does this have on personas' interactions with other students?

In the proposed system we hypothesize that Ayden will benefit the most from this change to the course structure as they have the most intrinsic motivation for collaboration with peers in learning the material. We expect that the peer interactions will have a different shape then observed previously, as students will no longer feel forced to crowd-source solutions to the peer-interview questions and trying to come up with solutions together. Instead, having tinkered through learning during the lab, they can now review the interview document with their lab partner and collaborate in mastering the material by reviewing one correct thought process for answering and **analyze** level question on the relevant topics.

4.4 Performance

1. What impact does this have on personas' ability to correctly assess their own performance?
2. What impact does this have on personas' performance?

Recall that real performance of the personas are as follows, Ethan **evaluates**, Ayden **applies**, and Umbrielle **understands**. Based on information collected in interviews, we theorize the personas reflective-assessment of their mastery as follows under the staff-interview model: Ethan **analyzes** while Ayden and Umbrielle **remember**. Here, we see that each student underestimates their ability. While Ethan and Umbrielle only underestimate their ability by one level, Ayden actually jumps two levels

down in their estimate of their own mastery of the material. This suggests that the staff-interview model of grading has a negative impact on student confidence as the student cannot accurately understand their performance in the course through the feedback provided on aggregate including in their staff-interview.

Based on the research detailed in the background section of this document and our understanding of the personas, we theorize that under the new peer-interview model, the following will hold true. Ethan will still under-assess their mastery level as **analyze** as this method does not resolve his imposter syndrome. Ayden will now correctly assess their own mastery level of the material as **apply**, as they receive more feedback from their peers throughout the semester, demonstrating one success in this method to allow students to accurately gauge their own level of mastery in the course. Finally, Umbrielle has the greatest potential to be impacted by this change in the course structure. To the best of our knowledge Umbrielle will either thrive or dive. If Umbrielle behaves ethically and completes the peer-interviews as assigned, she will now correctly estimate her level of mastery in the course as **apply** as her underlying persona is likely to change throughout the semester as she feels adequately resourced to complete the course with solutions to review and peers to assist in her learning. Alternatively, Umbrielle might continue to feel that she is not adequately resourced and continue to take unethical actions in the course. If so, then we expect that Umbrielle would fall to only a **remember** level of understanding.

4.5 Ethics

1. What impact does this have on personas' ethical behavior in the course?

Umbrielle and Ayden each behaved unethically under the old model. In the new model they would each feel more supported to learn the material. Umbrielle may still find it necessary to steal solutions for the lab from the internet, but in peer-interviews, she and her partner would be given interview questions and detailed hints and solutions that should allow them to learn the material deeper than they would have in the previous model as they no longer feel as much pressure to "get the right answer" in the staff-interview as formative assessment and instead can take the time to learn what they

can from the resources provided. So long as Umbrielle has enough time to learn from these interview documents, and she feels comfortable reaching out to the course staff when the interview documents are not sufficient to explain some topic, then she should be encouraged to behave ethically in the peer-interviews while still striving for her "A" in the course.

5 Future Work

This initial case study shows some promise to the value of peer-interviews, but leaves us with many more questions to answer.

Execution and Measurement: One next step in this research is to suggest how to measure the success of the proposed peer-interviews, then design and execute the experiment to collect and analyze data while summarizing the benefits and challenges of the model that should be considered prior to future iterations.

Staff Reflection Phase: One important advantage of the peer-interviews for the course staff is that the interview data is returned to course staff about one week earlier than it is with staff-interviews. Accordingly, if no students complete staff-interviews, then the course staff would be able to construct their action plan for course improvement based on the student reflections earlier and be able to deliver effective change to the classroom more rapidly. But what impacts would this have on the effectiveness of the review process if course staff had not actually completed an interview with a student and directly observed where students are struggling? Will the staff reflection phase still be as effective?

Why does it work: The current proposed method of peer-interviewing suggests value when implemented correctly, but why exactly does this work well for so many students? What aspects of the student learning environment exist in this modified course structure that could be leveraged in other aspects of the course? How would that change be implemented? When would it not be wise to make such a change? Does it work as well in lower division courses as it does in upper division courses?

Ethical guard-rails: Any person, with the right reason, will do something unethical. What guard-rails are important to apply to the course structure to better engage students in completing the lab interview reflections and deter students from lying about completing the interviews and submitting a falsified self-reflection?

6 Conclusion

We detailed a method of peer-to-peer reflective interviewing to engage students in a highly scale-able manner that improves student agency in learning. We have demonstrated how students from the Fall 2022 session CSCI 3155 might behave in the new model of interview grading. Specifically, we have demonstrated how these student personas would have an increased ability to assess their own learning and even improve their mastery of the course materials as we shift student focus away from "getting the right answer" and toward developing a deep understanding of why a correct answer is correct. Finally we detail some of the challenges that these personas will still face under the peer-interview model as they may choose to continue to behave unethically.

References

1. Zweben, S., Bizot, B.: 2015 taulbee survey continued booming undergraduate cs enrollment; doctoral degree production dips slightly Accessed: 2023-09-20.

2. Zweben, S., Bizot, B.: 2018 taulbee survey undergrad enrollment continues upward; doctoral degree production declines but doctoral enrollment rises Accessed: 2023-09-20.

3. Zweben, S., Bizot, B.: Cra 2022 taulbee survey: Record doctoral degree production; more increases in undergrad enrollment despite increased degree production Accessed: 2023-09-20.

4. Anderson, L.W., Krathwohl, D.R.: A taxonomy for learning, teaching, and assessing: A revision of Bloom’s taxonomy of educational objectives: complete edition. Addison Wesley Longman, Inc. (2001)

5. Flaherty, C.: When grading less is more Accessed: 2023-09-15.

6. Stommel, J.: How to ungrade Accessed: 2023-09-15.

7. Owens, K.: A beginner’s guide to standards based grading Accessed: 2023-09-06.

8. Chen, L., Grochow, J.A., Layer, R., Levet, M.: Experience report. In: Proceedings of the 27th ACM Conference on on Innovation and Technology in Computer Science Education Vol. 1, ACM (jul 2022)

9. Mittell, J.: Rethinking grading: An in-progress experiment Accessed: 2023-09-15.

10. Stommel, J.: Ungrading: A bibliography Accessed: 2023-09-15.

11. Shepard, L.A.: Ambitious teaching and equitable assessment Accessed: 2023-11-12.

12. East, J.P., Schafer, J.B.: In-person grading: An evaluative experiment. SIGCSE Bull. **37**(1) (feb 2005) 378–382

13. Ruehr, F., Orr, G.: Interactive program demonstration as a form of student program assessment. J. Comput. Sci. Coll. **18**(2) (dec 2002) 65–78

14. Grunwald, D., Boese, E., Hoenigman, R., Sayler, A., Stafford, J.: Personalized attention @ scale: Talk isn’t cheap, but it’s effective. In: Proceedings of the 46th ACM Technical Symposium on Computer Science Education. SIGCSE ’15, New York, NY, USA, Association for Computing Machinery (2015) 610–615

15. Martin, F., Ritzhaupt, A., Kumar, S., Budhrani, K.: Award-winning faculty online teaching practices: Course design, assessment and evaluation, and facilitation. *The Internet and Higher Education* **42** (2019) 34–43

16. Berge, Z.L.: Changing instructor’s roles in virtual worlds. *Quarterly Review of Distance Education* **9**(4) (2008) 407–414

17. Alam, A.: Employing adaptive learning and intelligent tutoring robots for virtual classrooms and smart campuses: Reforming education in the age of artificial intelligence. In Shaw, R.N., Das, S., Piuri, V., Bianchini, M., eds.: *Advanced Computing and Intelligent Technologies*, Singapore, Springer Nature Singapore (2022) 395–406

18. Gehringer, E.F.: Electronic peer review and peer grading in computer-science courses. *SIGCSE Bull.* **33**(1) (feb 2001) 139–143

19. Smith, A.J., Boyer, K.E., Forbes, J., Heckman, S., Mayer-Patel, K.: My digital hand: A tool for scaling up one-to-one peer teaching in support of computer science learning. In: *Proceedings of the 2017 ACM SIGCSE Technical Symposium on Computer Science Education. SIGCSE ’17*, New York, NY, USA, Association for Computing Machinery (2017) 549–554