

SPWI: Principles of Programming Languages @
Scale: The Value of Student Collaboration

Spencer Wilson

University of Colorado Boulder

[illegible]

Look for:

BEGIN_REVIEW

and

END_REVIEW

[illegible]

Abstract. So there are students, and they pay the school money, but why should they keep doing that if there is a higher value to coursera courses compared to their experience at this institution?

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 number of staff resources available per student decreases. With this student to staff interaction 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-to-peer interviewing in an ungraded and self-reflective model

for pair assessment on complex lab assignments for the Principles of Programming Languages.

We propose a system for measuring effectiveness of education based on student self-reported ability and weight this with their exam performance. 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. Finally, we report on the benefits suggested by the study for peer-to-peer interviewing compared to TA interviewing found in this experiment.

We explore this value of peer-to-peer interviews by four metrics:

1. What impact does this have on students' completion of the course?
2. What impact does this have on students' ability to correctly assess their own performance?
3. What impact does this have on student performance?
4. What impact does this have on student satisfaction with the course?

2 Background

2.1 Cost

What is the cost for students and is it worth it? This question alone could take up several pages... but we'll keep it brief.

In review of the university data we see that the average fee collected by students in the course is \$3369.25. With 270 students that's \$909,000+ with a conservative estimate of cost per credit hour. However, consider that the total pay to direct staff of the course in Spring 2024 is around 47,000 in direct pay. Of course, graduate student pay is more complicated than the face value paid and total compensation of the course staff is closer to \$93,000. These numbers may seem shocking that only a 10th of the funds paid by undergraduates are going to direct course support, but it is important to note that the student tuition fees are paying for much more than the direct course staff, they are paying for the buildings, academic advisors, technical tools and support,

and so much more. So what support does a student get for their 3k? if we did a straight line we'd say \$344 of direct support. 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, and office hours) while a large group of students do not leverage these resources.

per lecture that's around \$56 per lecture or recitation session if we count it that way.

Supposing the student meets the university standard of 3 hours of work per credit hour, a 4 credit hour class like 3155 will be 12 hours per week. Per hour of class, reading, and homework the student is paying \$18.72 and that is the thing that I think actually matters. They are paying for the full package at a 'reasonable' cost (by the measure that they are currently willing to pay it and it feels small). It's not only about coming to lecture, but it is also about having challenges in homework's that grow there experience and build on whatever they learn in lecture. But if that homework content is freely available online, then what are they paying for? They are paying for the social environment. They emotional and technical support on the topics as they attend those lectures, recitations, OH, and online forums such as piazza.

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. SPWI, is that really what I want? Could I just cut out the Bloom's taxonomy thing all together? would the paper stand if I only talk about performance and mastery on an x - ✓ + scale? I think it's worth trying it. I might have to go any just write it up and then see if I need to backtrack tot he blooms thing. If I go that route, I think I'll have to delete the bloom block rather than commenting it. I can always go grab it from the old file if needed.

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" [1]. 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, it is often a good goal for the course as a whole.

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 instead focus on providing substantive feedback to our students [2] [3] [4]. 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 [4] [5][6][7]. This concept can be leveraged effectively in interview grading to emphasis formative feedback over a course grade for the student, helping to move students toward intrinsic learning rewards over extrinsic ones [8].

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 [9] [10]. However, it can be done at scale, by offloading the effort to support staff such as graduate teaching assistants and

graders [11]. 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" [2]. 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 and use 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 in the classroom, and the world of online learning [12] [13] [14]. 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[12]. 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 [15]. This method has been employed extensively in the online learning environment where scale is potentially 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 [12] [16]. 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. Participation: a formative assessment in which students **analyze** information through discussions during class sessions.
2. Labs: a formative assessment in which students **analyze** topics of interest and 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.

3. Grading Interviews: a formative assessment in which students **evaluate** their mastery on the lab material with twelve minute one on one interviews with 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.
4. Exams: the summative assessment in which students **create** novel solutions to relevant problems in a timed assessment that is manually graded by the course staff and returned to students with some limited qualitative feedback.

3 Experiment

[illegible]

BEGIN_REVIEW

[illegible]

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, peer interactions, and Conversational Ai such as ChatGPT. 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:

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. If 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.

[illegible]

END_REVIEW (find next BEGIN_REVIEW)

[illegible]

3.2 Proposal

In this experiment we propose one core change to the course syllabus. Here, interviews are not graded based on the students' correct answers to the interview questions, but instead purely on the students completion of the interview. 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 peer-to-peer interviews with self-reflective components that are reviewed and followed up on by their Teaching Assistant (TA) with interventions as necessary.

Each TA is as member of the course staff with an assured "**analyze**" level of learning on the material. In this course we had eight TAs comprised of two "**create**", three "**evaluate**", and three "**analyze**" level of course mastery.

The course is comprised of six labs which build off the knowledge of the previous lab. In the first lab, students perform both a peer-to-peer interview, then an additional interview with a teaching assistant. By the beginning of the second lab, students choose to either spend the semester in interviews with a TA, or in peer-to-peer interviews.

3.3 Interview Process

In each model of interviewing the 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 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 had large upfront cost, this sample of the grading process 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 [15].

Interview Phase Consider hypothetical students Ethan and Ayden have just completed lab three as a student team.

Peer to Peer Interview Ethan and Ayden select a time to meet in-person, or over zoom, and discuss what they learned during the lab. They then download the interview question set for the lab and complete the interview together as a team. They are encouraged to complete the interview within thirty minutes; however, this is at the discretion of the student team. If neither student is able to answer the question, they can reach out on the course discussion forum to seek additional information on the topic. Here, we see more peer-to-peer grading interactions, giving students more autonomy in their learning and freeing the course staff to dedicate time to supporting student learning in other ways.

Teaching Assistant Interview Ethan and Ayden each sign up for one on one interviews with a member of the course staff. They attend the interview without prior knowledge of the questions that will be asked, and perform the interview in a twenty-minute slot (twelve minutes for lab 1). At the end of the interview, the course staff, tells the student how they performed on each question in an "X, ✓-, ✓, ✓+" scale and work 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 is 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.

Reflection Phase

Student Reflection and Action Planning Regardless of the interview method used, Ethan and Ayden now meet to review their performance on the 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 meant to discuss these plans together to

encourage cross-pollination of ideas. Each student submits this via a survey form that allows for the aggregation of student data.

Staff Reflection and Action Plan Next, the course staff review the student performance from their hosted interviews and enter notes about the student performance into a survey form that allows for the aggregation of student data. The course staff then gathers as a whole to review all the data provided both by the TAs and by the students to identify what students are succeeding with, and where they are really missing the mark. Collectively they discuss how this data can inform a change to the course lecture process, using the stores of knowledge that students have today to assist in filling in those gaps as we move on to new topics. Here the staff also has an opportunity to discuss what common issues and successes were observed during the interviews. In practice, this required a two-hour meeting with the full course staff at the end of each lab, after the completion of the interview phase.

Action Phase In the action phase, the course staff executes on their plan for improving the course lectures based on common findings in students' gaps. In an attempt to increase transparency of the process and build our students as conspirators 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.

3.4 Enrollment

SPWI: rework, no longer accurate... In this experiment students' self-selected to TA interviews or peer to peer interviews for the semester. The total course enrollment at the beginning of the term was 300, of which 60 selected to interview with a member of the course staff and 240 students selected peer to peer interviews. By the end of the

term, 18 students in the TA interview model withdrew from the course and 12 students in the peer-to-peer interview model withdrew.

3.5 Performance

SPWI: are the test scores an accurate measure of true bloom level? Consider the ethical behaviour of students. I suppose that if I'm right, then the cheating will decrease and this will be more correct. Although the more weight on some thing, the more likely that students will have wondering eye syndrome during their exams. Also consider the mater of timed exams for coding vs no more timed interview.... I suppose the in-class exercises become that much more important. I still think those should be the spot exams....

In an attempt to measure the impact of peer-to-peer interviews on student learning, we ask students to reflect on their performance in the course in comparison to the Bloom's taxonomy. Prior to the midterm and final exam, each student is asked to rank their learning on individual course topics and the course as a whole.

Additionally, we normalize the students' summative assessment scores to the Bloom's taxonomy. In review of the final exam content and grade distributions we categorized the student score as Bloom levels as follows (visualized in figure 1):

1. "create": the second and third positive deviation from the higher distribution
2. "evaluate": the first deviation from the higher distribution
3. "analyze": all scores below the first negative deviation of the higher distribution and above the first positive deviation of the lower distribution
4. "apply": the first deviation of the lower distribution
5. "understand": the second negative deviation of lower distribution
6. "remember": the third negative deviation of the lower distribution

We then weight the student self-reported Bloom level against the normalized summative assessment scores to construct a suggested true Bloom level of cognition for each student.

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 logistic 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 beginning 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 where 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. 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 content of those solutions. 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)."

CONTEXT NOTE - to integrate elsewhere in document: The below hypothesis are grounded in one observational truth. **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 by through honest means.** In general, the students want to learn. Many of them also want to be "handed" some of the information necessary to learn. In the new peer interview model we will be effectively handing them step by step solutions to problems to solve that are largely representative of the learning objectives for any given assignment of lab. Accordingly it is the students choice whether they will use that as a tool to accelerate their learning and better prepare for the in-class activities and exams, or if the student will make some other use of the tool.

4.2 Case Study

Let us demonstrate how these personas would have completed and interview in Fall 2022 with a member of the course staff facilitating the interview. In lab 2 students completed an assignment on authoring an interpreter for a subsection of JavaScript and interview with the Teaching Assistants. In the 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. However, 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 see one challenge of the previous TA interviews in that they better 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.

The data sent in by students for the interview prior to the beginning of lab 3 then helps the course staff to identify a common inability to accurately parse expressions in the language. Accordingly, the staff can adapted lectures during lab 3 to further emphasize visual parsing skills, while talking about the new topic of inference rules. Then we expect, during the reflection phase of lab 3, to observe that students have an increased mastery of parsing. That information can be leveraged to construct an action plan for future semesters to improve the instruction methods on parsing during lab 2 to assist in further improving the student learning experience.

4.3 Enrollment

1. What impact does this have on students' completion of the course?

Regardless of the method of interviewing used Ethan, Ayden and Umbrielle are likely to complete the course. However consider the students that represent a **remember** level of cognition. This represents a non-trivial amount of students enrolled in the course that would be impacted heavily by the proposed change. In the peer to peer model we see a split in the the **remember** students. Some of them want to get a passing great without putting in the required effort and we see that these students ultimately do not participate in the peer interviews. We regularly receive reports from their lab partners that they are refusing to participate. By lab 5, we expect that many of these students withdraw form the course as they realize that they will not be able to pass the course as they have not been completing the required work. Alternatively, we have students in the **remember** phase that grow throughout the semester to the **understand** phase. This growth is attributed to the peer collaboration build into the course where students are excited to assist in each others learning. These students go on to complete the course as demonstrated in figure 2.

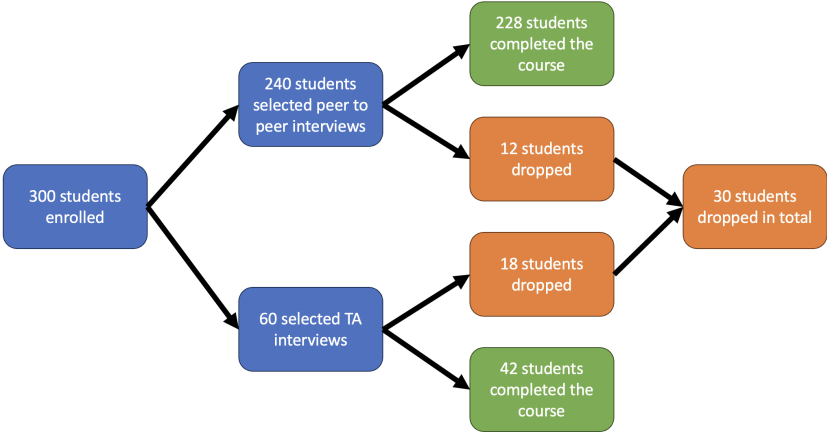


Fig. 2. The visualization is not updated, but something like this should be in the paper.

4.4 Performance

- 1. What impact does this have on students’ ability to correctly assess their own performance?
- 2. What impact does this have on student performance?

Recall that real performance of the student is as follows, Ethan **evaluates**, Ayden **applies**, and Umbrielle **understands**. Based on information collected in interviews, we theorize the persona personal assessment of their work as follows under the TA 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 current model of interviewing has a negative impact on student confidence. **SPWI: that’s not quite accurate... need to think more...**

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 to peer model for grading interviews 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**, demonstrating one success in this method to allow students to accurately gauge their own level of mastery in the course. Finally, Umbrielle will now correctly estimate their level of mastery in the course as **apply** as their underlying persona is likely to change throughout the semester. Umbrielle will see the greatest value in this change to course structure as she now has the resources that she feels is necessary to succeed in the course.

4.5 Satisfaction

- 1. What impact does this have on student satisfaction with the course?

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 to collabo-

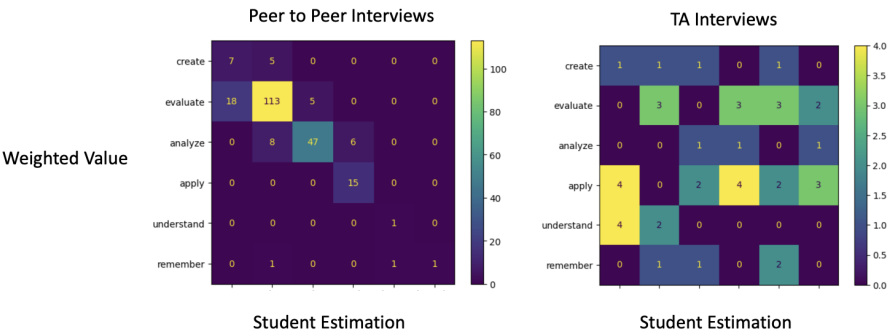


Fig. 3. SPWI: This is a great place for a graphic, but I'm unsure what an effective graphic will look like. I think this will be more apparent to me as I work on the presentation.

ration 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 interview questions and trying to come up with solutions together. Instead they can review the labs and solutions with their partner for the lab. Accordingly, we propose other mechanisms that encourage student collaboration on the course material.

QUESTION: Fall 16 and Spring 17 I ran review sessions after each lab as a course assistant. These semesters provided our weaker students with an ability to catch up on the lab content after the lab had finished. Can I use that information in this research as a personal record of events?

IDEA: We can leverage the CSPB videos in the interview documents to help students review the material. We would still offer them on day one and market them to students to improve performance throughout the lab and recommend selections for each day of lab work and tasks in the lab to attempt. We can adapt the videos on a case by case basis as improvements are warranted (driving updates by student performance and/or feedback for each semester). Note however, that there is a big difference in how topics are explained in the CSPB videos relative to how I - a near peer - explained the topics in AY17. Paripant 15 explains that the best part about working with a peer is that the discussion was devoid of all jargon.

[illegible]

END_REVIEW

[illegible]

5 Future Work

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

Timeline: One important advantage of the peer to peer interviews for the course staff is that the interview data is returned to course staff about one week earlier than it is with TA interviews. Accordingly, if the full class completed peer to peer interviews only, 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 TAs had not actually completed an interview with a student and directly observed where students are struggling? Would the staff reflection phase still be as effective?

Inclusion The collected data includes demographic data of students that may embed information about how students from traditionally marginalized and underrepresented communities are impacted by this course change. We are curious to see what information could be inferred from the existing data and consider further changes that better support these students.

Why does it work: The current proposed method of peer to peer interviewing has clear 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?

6 Conclusion

We have demonstrated the value proposition of using peer-to-peer interview grading over TA interview grading. This method leverages students themselves as a scale-able source of effective education actors in the learning environment with promising results at a time when student to staff interactions continue to decrease as course enrollment grows. This method, as implemented led to higher completion rates for students, a better ability for students to assess their own mastery of the material, higher mastery of the material, and higher overall satisfaction in the course. This model comes with an added benefit that course staff spends less time conducting interviews with students, allowing more time to review student performance and adapt teaching methods to meet the students' needs. While there is more work to be done, we hope that this structure continues to see adaptations that better enable our students' success.

References

1. 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)
2. Flaherty, C.: When grading less is more Accessed: 2023-09-15.
3. Stommel, J.: How to ungrade Accessed: 2023-09-15.
4. Owens, K.: A beginner's guide to standards based grading Accessed: 2023-09-06.
5. 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)
6. Mittell, J.: Rethinking grading: An in-progress experiment Accessed: 2023-09-15.
7. Stommel, J.: Ungrading: A bibliography Accessed: 2023-09-15.
8. Shepard, L.A.: Ambitious teaching and equitable assessment Accessed: 2023-11-12.
9. East, J.P., Schafer, J.B.: In-person grading: An evaluative experiment. SIGCSE Bull. **37**(1) (feb 2005) 378–382
10. Ruehr, F., Orr, G.: Interactive program demonstration as a form of student program assessment. J. Comput. Sci. Coll. **18**(2) (dec 2002) 65–78
11. 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
12. 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
13. Berge, Z.L.: Changing instructor's roles in virtual worlds. Quarterly Review of Distance Education **9**(4) (2008) 407–414
14. 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

15. Gehringer, E.F.: Electronic peer review and peer grading in computer-science
courses. SIGCSE Bull. **33**(1) (feb 2001) 139–143
16. 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