Using a Two-Stage Final Exam in an Intro CS Course: Student Perceptions and Grade Impacts

Celine Latulipe
celine.latulipe@umanitoba.ca
University of Manitoba
Canada

Lauren Himbeault
himbeaul@myumanitoba.ca
University of Manitoba
Canada

ABSTRACT

We present results from testing a two-stage exam format in a small, first-year programming class (n=24), including survey responses from students (n=8) about their experience with the exam format. Students reported liking the format due to a decrease in stress, helping them to better understand course concepts, and helping to improve their grades.

CCS CONCEPTS

• **Social and professional topics → Computing education**.

KEYWORDS

computer science education, assessment, two-stage exams

ACM Reference Format:
Celine Latulipe and Lauren Himbeault. 2023. Using a Two-Stage Final Exam in an Intro CS Course: Student Perceptions and Grade Impacts. In *Western Canadian Conference on Computing Education (WCCCE '23), May 4–5, 2023, Vancouver, BC, Canada.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3593342.3593355

1 INTRODUCTION

Two-stage exams are a relatively new format of high-stakes assessment in which students write an exam individually, and after the individual portion is complete and students submit their exams, they are then placed into groups and retake the same (or a very similar) exam as a group. Typically, students are given half as much time for the group portion as for the individual portion, since they should all be familiar with the questions. In this format, each student's final exam score (FS) is calculated as:

$FS = max(I * w + G * (1 - w), I)$

where I is the student's individual score, G is the group score, and w is the weighting factor (typically 85-90%).

This assessment format is purported to convey several benefits:

- It turns a summative exam into a learning opportunity, as students have the chance to discuss answers with others, realize their mistakes, and correct them. [9].
- It provides students a chance to debrief about the exam, which may reduce students' stress and anxiety levels [8].
- It gives students an opportunity to raise their grade slightly.

While two-stage exams have become increasingly popular in some STEM disciplines [6, 8], this format is less well-known in computer science (CS). Given that introductory CS courses exams often require students to solve programming problems by writing code, it is unclear how well this format integrates into evaluations and whether the shorter second stage is long enough for students to jointly complete programming problems.

As noted in prior research about conducting two-stage exams [7], these exams involve significant organizational overhead and require some special setup considerations:

- Everyone needs to take the same exam - question pools or multiple different exams are not suitable in this context.
- It is helpful to provide students with word puzzles or sudoku to do while they wait for the second stage to begin.
- The exam must be held in a separate room that can support moving people into groups for stage two. The exam cannot be held in a big gym with other concurrent exams, because stage two is loud and would be disruptive to others.
- There are a variety of ways to form the groups for the second stage [3], what is most critical is that students should not know ahead of time who is in their group.

Given the extra work involved in conducting a two-stage exam, it is worth investigating whether the effort is worthwhile in CS.

We used a two-stage format final exam in a very small introductory CS class (n=24) at the University of Manitoba in the Fall 2022 term. We report on the results in terms of grade impacts, as well as student perceptions gathered through a survey that was distributed after the semester was over and grades were submitted.

2 RELATED WORK

While two-stage exams are becoming common in other STEM disciplines [6, 8], there are only a handful of studies related to their use in post-secondary computing or IT education [2–4]. Belleville et al. created a novel, *inverted* two-stage test format for low-stakes quizzes in which students solved problems together first, and then were given similar problems to solve individually [1]. The purpose of this inversion was to promote student engagement through anticipation of transfer learning, and students preferred this format to individual quizzes followed by a TA walk-through.

3 TWO-STAGE EXAM DEPLOYMENT

In Fall 2022, a newly developed introductory programming class aimed at students with no prior programming exposure was offered at the University of Manitoba (n=24). This course was taught as a fully-flipped active learning class with students working on in-class activities throughout the semester in lightweight teams [5]. The final exam was worth 25% of students' final grade and was held in a 3-hour block scheduled by the registrar's office. The paper-based exam used the two-stage format, with the first two hours devoted to the individual exam and the last hour for the group portion.

There are different ways to form groups for two-stage exams [3]. We formed our groups by looking at student standing in the class and spreading both high performing and low-performing students across the groups in an attempt to balance knowledge and abilities.

4 RESULTS

We investigated how the two-stage exam impacted student grades and deployed a survey to students to gather feedback on their perceptions and experiences. For all but one of the 24 students in the class, the group grade was higher than the individual grade. For one high-performing student, the two grades were equal. For three students in the class, the two-stage exam helped them enough to push them up over a letter-grade boundary (e.g., from a C to a C+).

The survey, approved by our Research Ethics Board, was sent to students after the final grades were submitted and received eight responses (33% response rate). These students also provided consent to analyze and report on their detailed performance. For the eight students who participated in the survey, seven thought that the second stage exam improved their exam grade, one was not sure. The exam had a total of 50 points available and all 8 participants had final exam scores improved by the two-stage exam (mean difference = 9.6, min = 1, max = 19.5). Because the final exam was only worth 25% of the course grade, the average increase in final course scores for these eight students was only 0.48%, with a minimum bump of 0.05% and a maximum bump of 0.98%.

All participants stated that they would take another course that used the two-stage format, noting reasons such as the 'grade boost' (P4,P5), the 'fun experience' (P3) and leaving with a deeper understanding of the concepts (P7). All participants agreed that more instructors should use two-stage exams. One student suggested implementing two-stage assessments earlier in the term.

We asked participants whether they thought it made sense to use this format for exams with programming questions. All participants felt it made sense and P7 responded, "Yes, because of the complexity of some the long answer questions, answering it as a group to gain a better understanding can be beneficial." When asked about how the group collaborated on programming problems during the second stage, six participants said their teams 'jointly' created solutions by talking through the problems, one participant said they wrote their own solution, but sought input from their group, and P4 described their process this way: " ...each one of us would say how they did a particular question and then we would agree on which one made the most sense." We asked students to rate how much the second stage impacted their understanding of course concepts on a scale of 1 to 5, with 1 being a large decrease in understanding and 5 being a large increase in understanding. The mean rating was 4, the minimum was 3.

To assess impacts on student stress and anxiety, we asked the participants to comment on their feelings after completing the second stage of the final exam and five reported a sense of ease. They discussed feeling calmer and more confident because of the group conversations about the exam. Two participants reported some anxiety and nervousness from realizing what they did wrong in the individual stage, while one participant stated a benefit of the group stage was the ability to correct their mistakes.

5 DISCUSSION & LIMITATIONS

Because we employed lightweight teams to leverage social learning [5], incorporating team-based learning into the final exam made sense. The two-stage format appears to have had little impact on students' final course grades, resulting in letter grade bumps for very few students. However, the student experience of the format appears to be very positive, echoing results from studies in other STEM fields [6–8]. Students generally reported benefits from grade boosts, increased learning comprehension, and lower stress. Implementing two-stage assessments earlier in the term, as one participant suggested, would be a good way to help students get these benefits throughout the term.

We were unsure if the two-stage format would provide enough time for long format programming questions, but that did not appear to be an issue. We used a different group formation strategy than Cao et al. [3], and it is likely that the range of grade point differences we saw is due to creating groups with diverse abilities.

This is a small study with several limitations. All survey respondents received a C or better in the course - it is unclear how students receiving D or F grades felt about the two-stage exam. Our use of lightweight teams [5] likely made the two-stage exam format comfortable. Students in a lecture-based classroom, where there is no opportunity for social or collaborative learning, may find this format less comfortable.

6 CONCLUSION

We implemented a two-stage final exam in a small introductory programming course (n=24). All but one student had grades improved by the two-stage exam, but these resulted in a very small average increase in final course grades. Survey responses (n=8) indicate that all participants benefitted from the two-stage exam format, through easing student stress and enhancing understanding of course topics. Future work includes testing this exam format in classes that do not use team-based learning, and in larger classes.

REFERENCES

[1] Patrice Belleville, Steven A. Wolfman, Susanne Bradley, and Cinda Heeren. 2020. Inverted Two-Stage Exams for Prospective Learning: Using an Initial Group Stage to Incentivize Anticipation of Transfer. In *Proceedings of ACM SIGCSE 2020* (Portland, OR, USA) *(SIGCSE '20)*. ACM, New York, NY, USA, 720–738. https://doi.org/10.1145/3328778.3366938

[2] Yingjun Cao and Leo Porter. 2017. Evaluating Student Learning from Collaborative Group Tests in Introductory Computing. In *Proceedings of ACM SIGCSE 2017* (Seattle, Washington, USA) *(SIGCSE '17)*. ACM, New York, NY, USA, 99–104. https://doi.org/10.1145/3017680.3017729

[3] Yingjun Cao and Leo Porter. 2017. Impact of Performance Level and Group Composition on Student Learning during Collaborative Exams. In *Proceedings of ACM ITICSE 2017* (Bologna, Italy) *(ITiCSE '17)*. ACM, New York, NY, USA, 152–157. https://doi.org/10.1145/3059009.3059024

[4] Qianni Deng and Xinjian Luo. 2018. PipE2: An Innovative Pipelining Design for Collaborative Two-Stage Exams. In *Proceedings of ACM SIGITE* (Fort Lauderdale, Florida, USA) *(SIGITE '18)*. ACM, New York, NY, USA, 38–43. https://doi.org/10.1145/3241815.3241850

[5] Celine Latulipe, N Bruce Long, and Carlos E Seminario. 2015. Structuring flipped classes with lightweight teams and gamification. In *Proceedings of ACM SIGCSE 2015* (Kansas, USA). ACM, 392–397. https://doi.org/10.1145/2676723.2677240

[6] Clara L. Meaders and Yalila Vega. 2021. Collaborative Two-Stage Exams Benefit Students in a Biology Laboratory Course. *Journal of Microbiology & Biology Education* 0, 0 (2021), e00138–22. https://doi.org/10.1128/jmbe.00138-22

[7] Genevieve Newton, Rebecca Rajakaruna, Verena Kulak, William Albabish, Brett H Gilley, and Kerry Ritchie. 2019. Two-Stage (Collaborative) Testing in Science Teaching. *Journal of College Science Teaching* 48, 4 (2019), 64–73.

[8] Brian Rempel, Maria Dirks, and Elizabeth McGinitie. 2021. Two-Stage Testing Reduces Student-Perceived Exam Anxiety in Introductory Chemistry. *Journal of Chemical Education* 98 (07 2021). https://doi.org/10.1021/acs.jchemed.1c00219

[9] John F. Zipp. 2007. Learning by Exams: The Impact of Two-Stage Cooperative Tests. *Teaching Sociology* 35, 1 (2007), 62–76. https://doi.org/10.1177/0092055X0703500105