

Transitioning Two-Stage Exams to an Online Class

Jill A. Parsell¹, Allissa Blystone¹, Adrienne E. Williams^{2*}

¹Department of Biological Sciences, Orange Coast College, Costa Mesa, CA

²Department of Developmental and Cell Biology, University of California, Irvine, CA

Abstract

Students often find biology courses to be very difficult and isolating, particularly if they identify as part of a group that has been historically excluded from STEM. Some of this anxiety and isolation comes from high-stakes exams. We decided to use the collaborative structure of two-stage exams to try to overcome the isolation of assessment. In two-stage exams, students take an individual exam, and then immediately get into groups and take the exam again, discussing the questions and the rationale behind the answers. Their exam scores are a combination of the two attempts. Our move to emergency online learning because of the COVID-19 pandemic forced us to try our two-stage exams online. In this Teaching Tools and Strategies essay, we discuss our process of offering two-stage exams online at two different institutions: a two-year Community College and four-year Research University. We share feedback from the students and discuss our iterative improvements to two-stage exam use.

Citation: Parsell JA, Blystone A, Williams AE. 2022. Transitioning Two-Stage Exams to an Online Class. CourseSource 9. <https://doi.org/10.24918/cs.2022.29>

Editor: Justin Shaffer, Colorado School of Mines

Received: 9/14/2021; **Accepted:** 4/19/2022; **Published:** 10/3/2022

Copyright: © 2022 Parsell, Blystone, and Williams. This is an open-access article distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License, which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original author and source are credited.

Conflict of Interest and Funding Statement: None of the authors have a financial, personal, or professional conflict of interest related to this work.

Supporting Materials: S1. Online Two-Stage Exams – Sample Questions and Survey: Sample questions with percent correct, survey results

***Correspondence to:** 2014 Biological Sciences 3, University of California Irvine, CA 92697. Email: adriw@uci.edu

INTRODUCTION

Active learning is a collection of pedagogical techniques that has gained acceptance as a more effective way to cultivate and enhance learning in a classroom, even if the classes are large (1, 2). Examples of active learning in large lectures include small group problem solving, think-pair-share, and live classroom polling. We regularly utilize active learning in our face to face Anatomy courses, but we were still concerned about equity and attitudes in our courses. At the community college we were seeing high withdrawal rates from our Anatomy class, particularly among Latinx students. At the research university we saw high levels of anxiety during Anatomy exams. Because of this, we decided to add some version of two-stage exams to our courses with the intent of increasing the feeling of community and collaboration among the students, particularly around exams. With the forced switch to online courses during the COVID-19 pandemic, we decided to continue our plan to add online two-stage exams to increase student collaboration during this isolating time.

Student collaboration in real time in an online class seems to have benefits beyond that of collaboration in an asynchronous environment. Duncan and colleagues studied the synchronous and asynchronous messaging by MBA graduate students, and found that students who participated more synchronously performed better in the class (3). A 2014 study by O’Flaherty and Laws showed remote undergraduates in a first year biological science course had higher grades when working with tutors concurrently via web conferencing than when working with them asynchronously (4). A study by Wang and Wang separated pre-service science teachers into independent

work or synchronous work through meetings or discussion groups, with a face-to-face control (5). The researchers found that synchronous learners had higher social presence, higher cognitive presence, and demonstrated better teaching practices than asynchronous learners.

Two-stage exams utilize synchronous student collaboration, though typically in a face-to-face environment. In a “classic” two-stage exam, students complete a summative exam individually in a proctored environment. They then physically move into small groups and complete the same exam as a group, submitting a single group answer sheet. Their score is then some combination of both scores (6-8). Several studies have shown that student scores improve significantly on the group exam, for instance in veterinary physiology (9), nursing education (10), general medical education (11) and exercise physiology (12). While some studies find long-term learning benefits (10, 13), others find that there is little evidence that the improved learning is maintained by students to the end of the course (e.g., 14, 15). Our use of two-stage exams was not primarily chosen to increase long-term student performance (although that would of course be a benefit), but to create a clear course objective of community and collaboration among our students. This, plus other limitations described below, are why we chose to use formats that were different from the classic two-stage exams.

We were excited about adding collaboration (and hopefully learning) by using two-stage exams in our online classes and felt this was important to pursue in order to try to increase student self-efficacy and inclusiveness while reducing anxiety in students (16, 17). We were not able to find resources to help us shift the “classic” components of two-stage exams to an

online environment. After several discussions, we created and implemented collaborative exam plans for our online Anatomy courses, and we share those techniques with you here in hopes it will allow you to pursue similar techniques.

IMPLEMENTATION

We describe below how two different instructor groups in two different college settings decided how to offer group exams while teaching online. Both courses focused on Anatomy, but the format of the online group exams is relatively content-agnostic. For each section below, we will discuss how the issue was addressed in the Community College and Research University settings. It is highly likely that the reader will find elements of the solutions at both institutions to be applicable to their own setting. Pick and choose (and improve) as desired.

Students and Setting

Community College

At the community college, Anatomy with lab is taught as a lower-division biology elective. The majority of students take this course as a prerequisite for professional health programs, primarily nursing and physical therapy. In this course, about 41% are underrepresented minorities (primarily Latinx), and upwards of 68% are female. Out of 260 students each semester, many students are taking this course during their first year, if not first semester, in college. Due to work-life constraints for many, and therefore a lack of applicable time to study, approximately 35% of students withdraw from the course. Because of the necessary “A” or “B” required by professional schools, many also decide to withdraw and take the course again. This community college uses 16-week semesters.

Research University

At the research university, Anatomy with lab is taught as an upper-division biology elective. Almost all the students (about 150 per quarter) are in their last year and about to graduate. Many are pre-med and are taking the class out of interest. Others are nursing students or pre-health students who need the course as a prerequisite. At this university, about 47% are first generation, 29% are underrepresented minorities (primarily Latinx), and 65% are female. The university uses 10-week quarters. Students are not allowed to drop from this course once it has begun.

Existing Course Design

Community College

Anatomy meets two days a week for lecture for 85 minutes, with two 3-hour labs taught by either the lecturer or adjunct faculty. Each lecture covers a complete chapter, creating a fast-paced high-content setting. Students are provided detailed reading guides to complete prior to attending class. During class sessions, we utilize active learning strategies by facilitating practice questions with live polling, and questions and answers to provide feedback. During the COVID-19 pandemic, we provided pre-recorded lecture videos for students to access on their own schedule. Chapter exams were given weekly during class time. Student-driven question and answer sessions with the instructor preceded each exam. Exams were closed-book and were proctored by software within Canvas.

Research University

Anatomy meets three days per week for lecture for 50 minutes, with one weekly 3-hour lab taught by graduate teaching assistants and managed by the instructor. Our pace is similar to the community college: every day's lecture and every week's lab is a full chapter and anatomical system. We also assign reading guides for each chapter as the primary content delivery. Class time is reserved for small group problem solving and answering student questions. During the emergency switch to online learning, the notetaking remained the same, and no video “lectures” were provided. We offered online office hours during class time on Mondays and Wednesdays, and gave all four midterm exams on Fridays. Our institution strongly encouraged the instructors to make all exams open note (and thus open to online searching) and to not use proctoring software during COVID online courses.

Initial Decisions

The differences in institution, course and student populations led us to choose slightly different formats for the online two-stage exams. In “classic” face-to-face two-stage exams, students sit in a room together, take a proctored exam individually, then arrange themselves into groups and take the exam (or subset of the exam) again with group discussion and a single group submission. With a move online, exam submission options and student interactions are limited and access to the internet is more open. Below we describe the different two-stage exam formats we decided to use.

Community College

Our normal exams are closed-note; if students are allowed access to their notes or the internet, they can rather easily find the correct answers. We did not want to move too far away from our standard exam format, because the learning objectives for our anatomy course do require students to have a solid memorized list of anatomical structures and their functions in addition to some higher-order thinking. Our institution encouraged the use of proctoring software during online exams, so we elected to keep our two-stage exams also proctored. This meant we needed a separate discussion time, because students could not talk in breakout rooms during a proctored quiz. We therefore designed a discussion phase to the exam, utilizing Zoom breakout rooms, between the two individual exams. To reduce any student inclination to spend the discussion time searching online for the correct answers, we gave students a discussion “quiz” with only question stems provided. Students submitted responses that documented their group discussion. For the second quiz, instead of requiring a single group submission, students left Zoom and again took a proctored individual quiz.

Research University

At the research university we had the same limitation, that there was no way to submit a group quiz in Canvas, and a similar concern, that students would spend discussion time just looking up the correct answers. Additional constraints in our class were open-note exams, highly experienced upper-division BioSci majors (who are excellent Googlers), and tight class times of 50 minutes. To deal with the tight time window for exams, we decided to hold four two-stage “Group Quizzes” on Fridays separate from midterm Fridays. To match the community college faculty, we created a discussion time in between the two exams (individual Quiz 1, discussion time, individual Quiz 2). To address the “looking up answers during the discussion” issue,

the instructor gave students the Quiz 1 questions and answers to discuss, but made the Quiz 2 questions of the group quiz slightly different from the Quiz 1 questions.

Two-Stage Exam Formats

Since much of our administration of group quizzes (and online group quizzes in particular) was organizing timing and access, the details below are organized by a timeline. At both institutions, exams, two-stage exams and group quizzes were administered within Canvas. Examples of test questions are provided (S1. Online Two-Stage Exams – Sample Questions and Survey).

Community College

We chose three of the 15 weekly exams to integrate our two-stage exams. All exams were synchronous, including the two-stage exam format which had three specific assignments to complete within time limits. The two-stage exam took 80 minutes, as opposed to the 35 minutes of a regular weekly exam. See Table 1 for the timeline.

Research University

We chose five weeks during the quarter that did not have regularly scheduled midterms, and scheduled group quizzes for those Fridays. Students were told at the start of the term that exams were synchronous during scheduled class time on Fridays for the entire quarter, so all students had arranged their schedules to be present in the Zoom meeting. The group quizzes were tighter in focus than the midterms, only covering the material from that week (2-3 reading guides). The first instance was a practice group quiz to give students experience with the format without any points. The entire process took 30 minutes. See Table 1 for the timeline.

Scoring and Grading

Two different methods were used to calculate the scores associated with the two-stage exams. We provide the details as examples to the reader.

Community College

Each chapter exam is worth about 50 points, and since we drop the lowest exam scores, we wanted the two-stage exams to be worth the same number of points as the other exams. A regular exam had 25 questions worth 2 points each. For the two-stage exam, we made the 25 questions on the initial exam worth 1.8 points each (90% of the 50 points) and made the three questions on the second stage worth 1.67 points each (10% of the 50 points).

Research University

The group quizzes had four individual questions and four group questions. The scores on these two quizzes were summed, and if students answered four or more questions correctly of the eight, they received full credit for the activity. Students with fewer than four correct received partial credit. The updated credit was uploaded to Canvas as their “Group Quiz” score, which was worth 5% of the course grade. The lowest group quiz score was dropped.

STUDENT FEEDBACK

We collected different types of feedback from our students at the two different institutions, so we start with this feedback and then describe our modifications in the Iterations section.

Community College

During the first semester, students were loudly dissatisfied with the two-stage exam. Our discussions with students indicated that most students had never experienced this type of exam format, which may have led to their distrust of the process. Their chief concerns were as follows:

- Fairness: students complained that some people didn’t participate in the discussion, but could listen in and then “reap the benefits” on their second exam.
- Comfort: students didn’t like the inconsistent student membership in their breakout groups, which were randomly assigned each time. They didn’t know anyone in their groups and said they would have felt more willing to talk if the groups were the same every time.
- Group answers: Students worried that poor understanding by others in their group would affect their grade. Even after explaining the components of the exam, students could not appreciate that the second quiz was individual. If they differed in opinion from the group, they still had the liberty to answer differently on the quiz.

Research University

We asked our students about the online group quizzes via a survey at the end of the second quarter of implementation (IRB for exempt research was obtained via self-determination, as is standard at our institution). While 82% agreed that the group quizzes helped them understand course content and 78% said it increased their time studying anatomy, only 72% agreed that they made them feel connected to other students, and only 62% said group quizzes made them feel like they could be successful in the class. Overall, most students agreed (76%) that group quizzes should be retained when the class returns to face-to-face. The primary frustrations that these students reported were as follows:

- Several students reported that their groups had overly quiet members who didn’t participate in the discussion (groups were held constant for each exam)
- They found the (very tight) time restrictions for the quizzes to be stressful
- Only 73% said the Quiz 2 questions were similar to the Quiz 1 questions (even though the instructor thought they were similar).

ITERATIONS

Despite the differences between the implementations of two-stage exams at the two institutions, the frustrations felt by the students and the instructors ended up being rather similar. Below are the changes that we have made so far. Interested readers can see the positive result of these changes in the survey results from the research university in S1. Online Two-Stage Exams – Sample Questions and Survey.

Goal 1: Improve communication in discussion groups

In a discovery which will surprise none of our readers, students had strong opinions about their small groups at both institutions. Quiet students felt uncomfortable talking to strangers, and talkative students felt frustrated if they were trying to discuss the problems but other students were just listening. We at the community college saw a marked improvement in attitude during the second semester (Spring 2021) when we allowed students from the same lab sections to self-assign into groups and made the groups smaller (4 students) and constant. Moving forward, we will keep students from the same lab together, but we would also like to be sure each group that has Latinx students contains more than one, since we want to reduce the withdrawal rates among that population. We at the research university also more thoughtfully organized groups. We made groups smaller (4 students rather than 7) and regularly randomized groups during the term (the research university students preferred new groups each time). We provided better training on how to have an effective discussion, especially incorporating the importance of having each group member speak. In the future, we will assign a different facilitator for each group during each quiz with the job of encouraging all members to participate.

Goal 2: Better scaffold the biology discussions

The emergency move to online instruction meant we were learning how to design a new online summative exam format as well as this new two-stage exam format. Now that we have a better understanding of how well students answer online summative exam questions, we feel like we can write better questions for two-stage exams. At the community college, we plan to generate questions for the two-stage exam that will stretch students a bit more during the discussion time, and will also be difficult to find in their notes. We at the research university have already migrated to un-searchable questions, but want to improve students' ability to transfer their knowledge to slightly different versions of the questions. At the community college, we set up the discussion time with more scaffolding for students. We sometimes provided an image "hint" during the discussion time that was not provided during the initial exam. We also provided tips for solving the problem that also encouraged collaboration such as, "First, write down all the steps of synaptic transmission in order as a group." At the research university where questions on the second exam are different, we plan to guide students to prepare for an isomorphic question ("What are the primary functions of other brain regions not in these answer options? Make a short list in case the Quiz 2 question asks about different regions."). We have also learned how to make the Quiz 2 questions more similar to the Quiz 1 questions. Interestingly, in many cases students did not greatly improve on the second, isomorphic questions which means there is plenty of room for improvement in the discussion prompts. We hope that these changes will help the more tentative students feel like they can contribute and will also increase confidence and performance during the second exam.

Goal 3: Increase collaboration during the second quiz

With additional practice teaching online, we have become less concerned with cheating and more comfortable creating times that encourage students to work together and use the resources they have prepared. Both instructor groups decided that isolating and proctoring students during the second exam was not necessary, and in fact keeps anxiety high and collaboration low. We kept the middle discussion time because we think it

contains good biology-related conversation. But we now allow students to remain in their breakout rooms for the second test, and tell them they can use their notes and discuss the answer options while taking their individual quizzes. This is doable in both of our classes, since the point value of the activity is low and our goal for the two-stage exams is community rather than evaluation.

We were not primarily focused on student learning with the two-stage exams, although we of course would like for this to occur and for students to feel like collaboration increases understanding of the material. In our Supporting File (S1. Online Two-Stage Exams – Sample Questions and Survey), readers can see the pre and post correct response percentages for our sample questions. For the six community college questions, the students had the same question for the first and second stages. In four of the six questions, students improved after discussion, while two of the six showed no change. The research university questions were different in the individual and group quizzes. One topic showed clear improvement after discussion, one was mixed in different years, one showed no change, and one showed a decrease after discussion.

To conclude, we authors all felt like group quizzes were a valuable addition to our online courses, although we are still working to switch ourselves from an assessment mindset to a collaboration mindset. We continue to refine these activities and have used them now in online, hybrid and face-to-face classes and in both Anatomy and Introductory Biology. Not only have we been able to create a useful time of collaboration to our class, but the tool has become something the students really value (see survey results in S1. Online Two-Stage Exams – Sample Questions and Survey). We've also learned the importance of getting feedback from the students during the process. What we assumed would be problems were not (looking up correct answers), and what we didn't expect to be problems were (tight time windows, quiet group members). Working together, we can make our classes more useful and rewarding for both instructor and students.

SCIENTIFIC TEACHING THEMES**Active Learning**

Two-stage exams traditionally have a strong active learning component. After the initial individual exam, students work in groups to solve exam problems and submit them together. In our versions of the online two-stage exam, students worked in small group breakout sessions to discuss their answers to the first exam before taking the second exam on their own.

Assessment

Two-stage exams are inherently an assessment tool. When given online, instructors can see performance on the exams immediately after they close. Students are able to self-evaluate their learning a) during the discussion after the first exam, when they hear how their group mates addressed the same problems, and b) upon the release of the scores. S1. Online Two-Stage Exams – Sample Questions and Survey shows the improvement that students made on the post tests for sample questions. Note that scores on the second iteration of the questions were not always higher.

Inclusive Teaching

Two-stage exams can be inclusive if all students are given an opportunity to speak in their small groups. We recommend that instructors use an initial practice exam as an opportunity to talk about ideal behavior during the discussion time, encouraging respect for all participants, encouraging every student to offer either their answer to a question, or a confusion about that answer. A rotating member of the groups can be assigned a facilitator role to make sure everyone is included and treated well regardless of their confidence with the material.

SUPPORTING MATERIALS

- S1. Online Two-Stage Exams – Sample Questions and Survey: Sample questions with percent correct, survey results

ACKNOWLEDGMENTS

The authors wish to thank the Biology Education Intersegmental Collaborative (BEIC) for its organization of our faculty learning community of 2-year and 4-year partners, and Samantha Elliott for her mentorship. The Office of Institutional Effectiveness at Orange Coast College provided significant help with survey administration and data provision for this project.

REFERENCES

1. Hake RR. 1998. Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. *Am J Phys* 66(1):64-74. <https://doi.org/10.1119/1.18809>
2. Freeman S, Eddy SL, McDonough M, Smith MK, Okoroafor N, Jordt H, Wenderoth MP. 2014. Active learning increases student performance in science, engineering, and mathematics. *PNAS* 111(23):8410-8415. <https://doi.org/10.1073/pnas.1319030111>
3. Duncan K, Kenworthy A, McNamara R. 2012. The effect of synchronous and asynchronous participation on students' performance in online accounting courses. *Accounting Educ* 21(4):431-449. <https://doi.org/10.1080/09639284.2012.673387>
4. O'Flaherty, JA, Laws, TA. 2014. Nursing student's evaluation of a virtual classroom experience in support of their learning Bioscience. *Nurse Educ in Prac* 14(6):654-659. <https://doi.org/10.1016/j.nepr.2014.07.004>
5. Wang J, Wang Y. 2021. Compare synchronous and asynchronous online instruction for science teacher preparation. *J Sci Teacher Educ* 32(3):265-285. <https://doi.org/10.1080/1046560X.2020.1817652>
6. Stearns SA. 1996. Collaborative exams as learning tools. *Coll Teach* 44(3):111-112. <https://doi.org/10.1080/87567555.1996.9925564>
7. Wieman CE, Rieger GW, Heiner CE. 2014. Physics exams that promote collaborative learning. *The Phys Teacher* 52(1):51-53. <https://doi.org/10.1119/1.4849159>
8. Shaffer JF. 2020. Student performance on and perceptions of collaborative two-stage exams in a material and energy balances course. *Chem Eng Educ* 54(2):52-58.
9. Giuliadori MJ, Lujan HL, DiCarlo SE. 2008. Collaborative group testing benefits high-and low-performing students. *Adv Physiol Educ* 32(4):274-278. <https://doi.org/10.1152/advan.00101.2007>
10. Rivaz M, Momennasab M, Shokrollahi P. 2015. Effect of collaborative testing on learning and retention of course content in nursing students. *J Adv Med Educ Prof* 3(4):178-182.
11. Lindsley JE, Morton DA, Pippitt K, Lamb S, Colbert-Getz JM. 2016. The two-stage examination: a method to assess individual competence and collaborative problem solving in medical students. *Acad Med* 91(10):1384-1387. doi: 10.1097/ACM.0000000000001185
12. Cortright RN, Collins HL, Rodenbaugh DW, DiCarlo SE. 2003. Student retention of course content is improved by collaborative-group testing. *Adv Physiol Educ* 27(3):102-108. <https://doi.org/10.1152/advan.00041.2002>
13. Cooke JE, Weir L, Clarkston B. 2019. Retention following two-stage collaborative exams depends on timing and student performance. *CBE Life Sci Educ* 18(2):ar12. <https://doi.org/10.1187/cbe.17-07-0137>
14. Kinnear G. 2021. Two-stage collaborative exams have little impact on subsequent exam performance in undergraduate mathematics. *Int J Res Undergrad Math Educ* 7(1):33-60. <https://doi.org/10.1007/s40753-020-00121-w>
15. Leight H, Saunders C, Calkins R, Withers M. 2012. Collaborative testing improves performance but not content retention in a large-enrollment introductory biology class. *CBE Life Sci Educ* 11(4):392-401. <https://doi.org/10.1187/cbe.12-04-0048>
16. Dewsbury B, Brame CJ. 2019. Inclusive teaching. *CBE Life Sci Educ* 18(2):fe2. <https://doi.org/10.1187/cbe.19-01-0021>
17. Cooper KM, Schinske JN, Tanner KD. 2021. Reconsidering the share of a think-pair-share: Emerging limitations, alternatives, and opportunities for research. *CBE Life Sci Educ* 20(1):fe1. <https://doi.org/10.1187/cbe.20-08-0200>

Table 1. Timelines for administering two-stage exams at the two institutions. See the text for iterative changes to this initial format after student feedback.

Time	Activity
Community College	
0:00	The weekly exam opens. The exam has an average of 30 multiple choice questions. The exam is proctored within Canvas with the camera feature required.
0:35	The weekly exam closes. When the exam closes, students do not get to see their score or the answers. Students enter the class Zoom session with their cameras on.
0:40	Students are assigned to a breakout room with 4-6 students.
0:41	The discussion participation quiz opens. Students individually access the quiz, which includes five questions from the individual exam that they have completed, and an additional question asks them to provide the student names in their group. The question stems are given, but no answer choices are provided, and students are asked to type a response that reflects the discussion of their group. The instructor and one course assistant visit breakout rooms to ensure cameras are on and students are on task. As soon as groups submit their participation quiz, they are instructed to return to the main session and wait with their cameras on for the duration of the assignment.
1:05	The discussion participation quiz ends, and students re-enter the main Zoom session if they haven't done so already. After all students have returned, the students are dismissed from Zoom.
1:10	The post quiz opens. Students take the proctored second quiz independently, with the same five questions from the weekly exam and participation quiz, now with answer choices provided. They are seeing these questions for the third time.
1:20	The post quiz closes.
Research University	
0:00	Students arrive in the main Zoom session with the instructor.
0:05	Quiz 1 opens. The quiz has four difficult multiple choice questions, and students are given only six minutes to complete the quiz (similar to the time constraints for the midterms and final). Students need to be present in the Zoom room, but cameras are optional. All students receive the same questions. When the quiz closes, students do not get to see their score or the answers.
0:11	Quiz 1 closes. We thank the students for their focus on the first quiz and then send them to Zoom breakout rooms for 10 minutes. We randomly sort students into 25 rooms of seven students, and reuse the same preassigned groups for all group quizzes during the quarter (see the Article's Discussion for whether this was a good idea).
0:12	The discussion quiz opens. Students receive a copy of the four quiz questions in a Canvas practice quiz that lists the questions as text (they are not answerable). The purpose of the discussion "quiz" is to give students access to the questions for discussion - it was not worth points. The instructor visits breakout rooms (in our experience, students do talk about the questions). We encourage cameras to be on but do not require it. The practice quiz closes after 10 minutes, and students begin to trickle back into the main room.
0:22	We close the breakout rooms and all students return to the main room.
0:24	Quiz 2 automatically opens in Canvas, and students begin taking it individually. There are 4 isomorphic questions that are intended to test similar anatomical concepts in the same formats as the original 4 questions.
0:30	The quiz automatically closes after 6 minutes. Students are free to leave when they finish, or stay to talk to the instructor.