Journal of Teaching and Learning with Technology, Vol. Vol. 13, Special Issue, pp.33-48.
doi: 10.14434/jotlt.v13i1.38519

Gradescope in Large Lecture Classes: A Case Study at Indiana University

Carrie A. Hansel
Indiana University, Bloomington
cahansel@iu.edu

Joshua D. Quick
Indiana University, Bloomington

Cathrine E. Reck
Indiana University, Bloomington

Anna Hungerfield Greene
Indiana University, Bloomington

Maggie Ricci
Indiana University, Bloomington

Anne Ottenbreit-Leftwich
Indiana University, Bloomington

Abstract: This multicase study explores how instructors at one higher education institution used an emerging technology, Gradescope, in their large classes to support the assessment of student learning. The study examined the practices of two instructors who taught large lecture classes with more than 100 students and used Gradescope to assist with the assessment of student learning. Results indicate that using Gradescope helped the instructors overcome specific challenges associated with large lecture courses, such as grading a large volume of assignments, providing timely feedback, and maintaining grading consistency across large numbers of students and graders. Course drop/fail/withdraw rates decreased and average course grades increased during the semester in which the instructors used Gradescope. Instructors reported increased grading efficiency and improved student satisfaction. However, the use of Gradescope was not entirely responsible for the positive impacts. Instructors also described adapting their teaching methods and assessment strategies to integrate Gradescope effectively.

Keywords: student response systems, large class, educational technology, active learning, student engagement

Large classes can offer a cost-effective and successful educational experience (e.g., Branca & Slusser, 2022). However, research has also suggested that large lecture classes can pose significant challenges for students (Kokkelenberg et al., 2008; Mulryan-Kyne, 2010). For instance, the sheer number of students can lead to feelings of disconnectedness for both students and faculty and decreased student engagement, and delays in providing timely feedback can lead to decreased student motivation. Reactions from both faculty and students indicate that traditional large lecture classes hinder student success. Without additional support, outcomes in these courses can vary widely and do not accommodate different teaching styles. To address these issues, faculty need to select instructional strategies and new technologies to address these shortcomings.

Both faculty and students struggle with feeling disconnected in large classes (Cash et al., 2017). Students have described large classrooms as "impersonal and anonymous…where neither peers nor

their instructor noticed whether they were absent or attentive" in contrast to small classes that "emphasized personal learning experiences…sense of community…with group or individual presentations and an emphasis on practical skill development" (Cash et al., 2017, p. 6). This feeling of being lost in the crowd can reduce students' motivation in the class and their subsequent course completion (Cash et al., 2017; Ng & Newpher, 2021). Students' motivation or participation (Lee et al., 2015) is critical to students' success and increased resilience in learning (Hartnett, 2012).

One method for addressing student motivation and participation, especially in larger courses, is timely, constructive feedback (Jacobi, 2018; Lim & Kim, 2002-2003), which has been shown to create needed connections and increase students' continuance in a course (Cash et al., 2017; Jensen et al., 2021). This is especially true when students need motivation to continue learning (Visser et al., 2002). Furthermore, feedback has been shown to positively impact students' deep learning (Higgins et al., 2002), skill improvement, and performance in the classroom (Bernius et al., 2022). However, many large-class faculty struggle to provide the timely and necessary student feedback at this large a scale (Cash et al., 2017; Lim & Kim, 2002-2003). Some faculty tackle this challenge of disconnectedness by creating smaller group activities to provide connection opportunities and a more personal learning experience (Cash et al., 2017).

Another challenge both large and small classrooms face is motivating students to engage with course content. Incorporating active learning increases student engagement by encouraging students to process course information through hands-on activities (Bolliger & Des Armier, 2013). Active learning has been shown to increase students' connection with the class and course satisfaction (Bolliger & Des Armier, 2013). Some have achieved active learning by utilizing a flipped classroom model. In a flipped classroom, students are provided with readings and multimedia materials to review prior to class and are asked to apply what they have learned to activities during class (Betti et al., 2022). This technique is effective with both large and small classes as well as when students are broken into small groups, where they can share highlights with the other groups about what they learned during the activity.

For example, Morosan et al. (2017) redesigned their course using a flipped classroom model. They provided voice-over slides presenting the content to 230 undergraduate students prior to class and then had students use mobile devices to engage in the classroom activities. The study found that this model improved students' technology and teamwork skills. Students reported they found the mobile device activities engaging but that sometimes they were distracted by other content on their phones.

Some flipped classroom model studies have shown positive effects on students' grades and improved student engagement but others have reported negative student attitudes or no difference from the traditional classroom experience (e.g., Betti et al., 2022). Students' attitudes toward learning are critical to consider when examining course design, as whether positive or negative, attitudes have been shown to impact academic performance (Nja et al., 2022). Positive attitudes have been linked with increased academic performance and negative attitudes with decreased performance (Nja et al., 2022). Additionally, research literature suggests that flipped classrooms positively influence students' attitudes toward learning, which increases students' academic performance (Ng & Newpher, 2021).

Regardless of discipline or classroom size, student engagement in active and collaborative instructional environments leads to improved student learning outcomes (see Kuh et al., 2005, 2007; Pascarella & Terenzini, 2005), such as increased motivation and course continuance (Cash et al., 2017; Jensen et al., 2021). Therefore, this is something instructors need to ensure they are examining when designing their courses. Lack of motivation stands as a primary reason for a high online student dropout rate (Hartnett, 2012; Visser et al., 2002) of up to 50% (Lee et al., 2015). Many factors contribute to this decreased learner motivation, such as feeling isolated and impoverished classroom interactions (Keller & Suzuki, 2004). To combat these factors, faculty need to know the learner

audience to select the most appropriate motivational strategy (Clayton et al., 2010; Politis & Politis, 2016) and a complementary technology to support student learning (Bouchard, 2009).

Student Engagement Technologies

Higher education's use of student engagement technologies began as early as the 1960s (Aljaloud et al., 2015; Habel & Stubbs, 2014). Such technologies provide a mechanism that taps into three facets of student engagement (Hollister et al., 2022): behavioral (active participation), cognitive (understanding), and affective (positive emotions). Student response systems (SRSs) focus on student engagement incorporated as part of the faculty's lecturing. After the faculty teaches on a concept, they pause to ask students to answer a prepared question that fits with the topic being taught (Gok, 2011), which has been described as "active adult learning" (Tuma, 2021). This process provides feedback to the students while also allowing the faculty to alter their future instruction to address errors identified from the individual and combined students' responses (Gok, 2011).

Using in-class technologies such as SRSs benefits and challenges both students and faculty. For the student experience, SRSs have been touted as being mostly a positive influence on student attendance, engagement, participation, and interaction with faculty (Aljaloud et al., 2015; Gok, 2011; Habel & Stubbs, 2014). Additionally, SRSs have resulted in increased student understanding of course material and improved student performance (Aljaloud et al., 2015). For faculty, SRSs have resulted in improved use of teaching and feedback strategies (Aljaloud et al., 2015), less grading time, and a better understanding of the concepts with which students are struggling (Gok, 2011). The challenges of using SRSs include wasted class time for both the faculty members and students due to technology issues, more student guessing, and increased faculty preparation time (Aljaloud et al., 2015).

Using SRSs in Large Classes

The evolution of digital tools can offer novel ways for faculty to give learners quality feedback (Jensen et al., 2021) while managing the increasing number of students within a course (Bernius et al., 2022). Stoerger and Kreiger (2016) used an SRS (i.e., "clickers") to add more student interaction and provide feedback on learning. Although student assignment scores increased, the faculty mentioned that it was challenging to get students to engage in large lecture courses differently. Educators have also been experimenting with automated essay grading (Bernius et al., 2022). However, this requires building a perfect answer and then comparing students' essays to the perfect answer. This can minimize creativity and multiple pathway solutions (Bernius et al., 2022).

McKnelly et al. (2023) outlined another example in a large lecture chemistry course that implemented a specifications grading system, which outlined preset criteria on what constituted satisfactory achievement for students to receive a certain grade. Students could choose which grade they wanted to work toward with the higher grades, an A or B, requiring students to demonstrate more mastery (McKnelly et al., 2023). They found higher overall grades at the end of the semester, and although during the semester students did not like the new approach, at the end of the course, their attitudes were more positive (McKnelly et al., 2023). Although teaching assistants preferred the new grading rubrics, students complained that they were more challenging to understand (McKnelly et al., 2023).

Gradescope

One digital technology, Gradescope, addresses the challenge of providing timely feedback to students in the large classroom. Created by Turnitin (https://www.gradescope.com/), Gradescope provides

faculty with options to flexibly create a variety of assignments, such as practice problems and tests. For example, faculty can use already existing paper-based activities or exams that students complete and then scan. Additionally, the faculty member can develop in-class digital activities that live inside the Gradescope app; students use their phone to type in the answers.

According to Turnitin, Gradescope allows students to easily submit assignments that require uploading pdfs, photos, and code from GitHub and Bitbucket while also providing faculty and their teaching assistants ease of grading within the digital platform. For paper-based exams or bubble sheets, faculty can scan the completed student assignments. This digitizes handwritten answers for use on the grading platform, where a rubric can be added to provide consistent grading among multiple graders.

Frequently used feedback can be saved on the platform, making it easy to add tailored feedback according to the students' needs while streamlining the faculty's grading time. The platform provides analytics on specific questions or rubric items, allowing faculty to see how their students are performing, and details on the wrong answers given by students.

Gradescope also fills in feature gaps within the learning management system (LMS), in our case, Canvas. For example, the Canvas LMS does not offer faculty the ability to grade all students' responses to each individual question on an assignment; whereas Gradescope does, which allows faculty to grade more consistently on a specific question across students. Another feature automatically divides work across multiple graders and permits parallel grading on "the same assignment and even the same question" (Turnitin, 2024a), which is not available in Canvas. Grades can be sent directly to the grade book or the students.

Aim of the Current Study

Given the potential impact Gradescope can have for large lecture classes, we sought to understand its use in the context of teachers' practices in large lecture classes. Specifically, we addressed the following questions:

1. What need did Gradescope meet in fulfilling faculty members' assessment aims?
2. Did Gradescope use correspond to any change in outcomes in the observed classes?

Methods

This multicase study (Merriam & Tisdell, 2016) aimed to explore how faculty in one higher education institution used Gradescope as an emerging technology in their large classes to support assessment of learning. Participating faculty members used Gradescope during the semester to assist with assessment of student learning.

Multiple data sources were collected for this research study, which allowed for triangulation that supports a more rigorous analysis (Yin, 2016, 2018). Each faculty member completed a questionnaire, provided the assignment descriptions, and consented to participate in a 60- to 90-min semistructured interview with the first author. Additionally, the faculty members granted access to anonymized grades for the individual assignments and final grades from spring 2023 through spring 2024. The study was approved by the university institutional review board (IRB).

Sampling Plan

We purposively recruited faculty in classes with 100 or more enrolled students who were listed as users of Gradescope, which provides the affordance of capturing richer information (Yin, 2016). Email

invitations were sent explaining the study. The sample size for this multicase study was limited to *N* = 2 to ensure that the complexity would be represented while preventing unwieldy data.

Participants

Table 1 displays the case profiles (pseudonym, number of courses taught using Gradescope, the number and type of course elements for which Gradescope was used, and the course names).

Table 1. Case profiles of courses.

Faculty name[a]	Role	Discipline	No. of courses or sections	Gradescope Use	Course name
Zoe	Course supervisor, coinstructor, and designer	Mathematics	12 sections	24 activities	Business Statistics[b]
Andrea	Course instructor and designer	Chemistry	3 courses	6 quizzes and 5 exams; 7 quizzes and 4 exams	Organic Chemistry I[b, c] Fundamentals of Chemistry and Biochemistry II[d]

Note. [a] Pseudonyms are used. [b] Spring 2023. [c] Spring 2024. [d] Fall 2023.

Zoe has over 20 years of experience as a mathematics professor. As the course supervisor for 12 sections of Business Statistics courses, Zoe decided to use Gradescope for 24 in-class activities. She did have other assignments that were not included in this analysis that were housed within Canvas and for which she did not use Gradescope (e.g., three exams, 24 homework assignments, and 103 formative assessments). Zoe also cotaught four sections of this course, while supervising the other eight sections. To maintain consistency across sections, Zoe oversaw making decisions on the course design for all sections. We elected to include Zoe's voice as representative of this case since she led the instructional design choices and supervised other instructors' uses of Gradescope.

Andrea has over 20 years of experience as a chemistry professor. She decided to incorporate Gradescope into her quizzes and exams for her spring 2023 and 2024 Organic Chemistry classes and her fall 2023 Fundamentals of Chemistry and Biochemistry II course. Each of Andrea's classes had six or seven quizzes and four or five exams. She did have other assignments that were not included in this analysis that were housed within Canvas and did not use Gradescope (e.g., 10 or 11 homework assignments, depending on the course).

Combined, these two faculty used Gradescope (either directly in their classes or in a supervisory perspective) for a total of 3,211 student enrollments. Only one student appeared in both faculty members' courses. However, the focus of the current study was on the faculty members' experiences. This is, however, a potential future point of investigation on the effects of Gradescope on student experiences.

Data Sources

Participants completed a demographic questionnaire and an individual, 60- to 90-min interview to gather more details about their rationale for using Gradescope as well as their overall experience with Gradescope. Specifically, these interviews focused on the faculty's aims and perceptions of outcomes in using Gradescope. The interviews were recorded using Microsoft Teams, and Microsoft Copilot generated initial transcripts. Deidentified Gradescope student scores on assignments/exams, in addition to final grades, were collected to compare student outcomes between semesters when Gradescope was and was not used.

Data Analysis

Once the questionnaires and semistructured interviews were completed, the automatically generated transcripts were pulled from Microsoft Teams. The first author reviewed and edited the accuracy of the transcripts, which was part of the interpretative analysis process (Johnson, 2011; Paulus et al., 2014). Multiple readings allowed the researcher to familiarize herself with the data. Thematic analysis (Braun & Clark, 2006) was used to identify recurring themes of faculty members' experiences with Gradescope.

Data Extraction

All data were extracted from internal institutional records and through the Unizin Data Platform (UDP). The UDP provides integrated, cleaned, and aggregated data on teaching and learning tools, such as Canvas. Institutional records were accessed to compile the official course grades of each student enrolled in the selected classes of each faculty. The UDP was used to compile assessment scores of Gradescope-associated submission records (e.g., the activities in Zoe's class and the quizzes and exams in Andrea's course). It was also used to count instances of Gradescope learning tool interoperability launches from the faculty's courses. Data pulls focused on the specified cases and the most recent iteration of the course before Gradescope was implemented in these courses. In Zoe's case, she conducted the non-Gradescope class activities through paper-based worksheets that were graded and then submitted into the Canvas system.

Institutional records were accessed for these two faculty to identify prior implementations of their Gradescope courses. The two faculty previously taught a specified course within the recent, 2-year period. These prior course outcomes and activities were used for comparison with the Gradescope courses. Andrea had previously implemented Organic Chemistry courses, but her Gradescope implementation was the first time she taught Fundamentals of Chemistry & Biochemistry II. Zoe had previously taught Business Statistics. Assessment activities measuring similar outcomes and using similar practices were compared for each faculty member's Gradescope and non-Gradescope implementations. This resulted in a total of 16 classes using Gradescope and four prior course implementations for comparison.

Descriptive Summaries

Summary statistics were computed for each class implementation. Specifically, the following measures were calculated: (a) class averages of the overall Canvas course score; (b) DFW outcomes; (c) average submission scores for specified assignments; (d) Gradescope launches from Canvas by faculty and students. Mean of group (i.e., class) means was calculated to identify general trends within the overall course and similar activities between the Gradescope and non-Gradescope implementations (if

available). As many of these implementations were single or "low *n*" occurrences, dispersion statistics were not computed. These measures provided comparison points of student outcomes between Gradescope and non-Gradescope course implementations.

Results

We observed generally positive perceptions of Gradescope by Zoe and Andrea. Additionally, we noted positive improvements in graded outcomes when comparing faculty members' use of Gradescope to previous versions of their course.

Faculty 1: Zoe

Zoe participated in the university's spring 2023 Gradescope pilot with her Business Statistics classes with 1,378 students enrolled in 12 sections ranging between 50 and 100 students. Although she used automated grading functions in Canvas for her quizzes and exams, she sought a solution for more efficiently and effectively assessing her students' learning through in-class, paper-based activities. These activities focused on core learning outcomes for business statistics and involved students answering questions or performing analyses on sample data sets. Prior to launching the tool, Zoe met with a teaching center representative to learn how to use the tool but said she generally "just figured it out."

In previous semesters, Zoe described three challenges to providing timely feedback for in-class, active learning activities: (a) grading the high volume of paper-based activities, (b) needing an electronic strategy for students at a distance due to religious holidays that spanned multiple weeks, and (c) ensuring only limited grading access to undergraduate teaching assistants (UTAs). In her interview, Zoe emphasized the importance of including active learning principles in her large lecture courses. According to the course materials, her course included 24 paper-based activities that students completed during class time, which typically took 20–30 min. Graduate teaching assistants (GTAs) would pick up assignments at the end of class for grading. Once the grades were entered into the LMS, the GTAs handed the papers back to students during another class period.

However, she encountered additional challenges with this model when the department switched from having GTAs to UTAs. Zoe was concerned about providing UTAs access to their peers' grades: "[I wanted them to have] access to only part of the [LMS'] gradebook…to be able to grade activities but not to have access to their colleagues' exam grades or homework grades or other grades." According to Zoe, the intensity of paper grading was challenging because of (a) handling the number of papers (which "was not sustainable because of being turned in on paper and the papers getting lost") and (b) complicated grading processes ("a grader had up to 20 parts to grade, which was overwhelming…when we were trying to teach in six weeks"). Zoe shared that in light of these challenges, a colleague suggested Gradescope to address the challenge of assessing the in-class group activities: "the problem was I wanted in-class group activities electronically graded—easily graded, easily returned…. [Gradescope could] facilitate the in-class activities because we weren't going to be able to keep up with it if it was on paper."

In her Gradescope classes' group work assignments, students would complete the 24 paper-based activities during class as usual but then completed the additional step of capturing a pdf image of the activity using the Gradescope app (see Figure 1). Zoe shared that students experienced some initial frustration as with any new technology, but after a while, "it's a relatively seamless tech thing." Students' complaints focused on having to do "the activities because it forces them to come to class and they don't want to," said Zoe. Students did not report any complaint about Gradescope, just as they did not "mention Canvas in their comments…[which] is a good thing," Zoe stated. Gradescope

provided a mechanism for Zoe to assign activity grading to her UTAs without providing access to the other grades within the LMS' grade book. In addition to privacy concerns, Zoe described that this virtual process of grading was "more secure and FERPA [Family Educational Rights and Privacy Act] compliant."

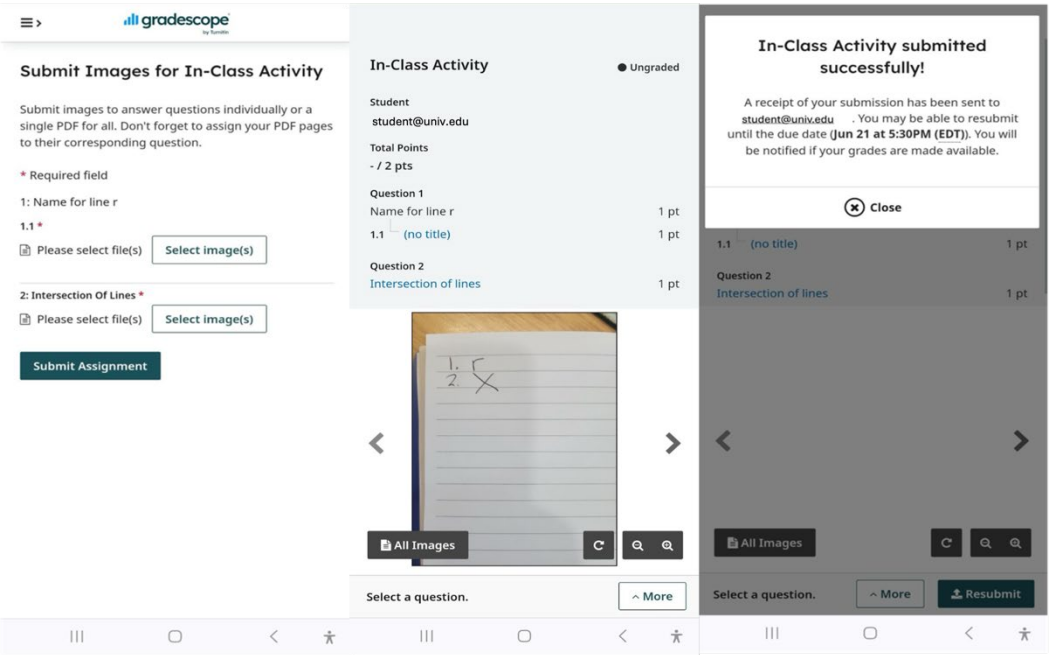

Figure 1. Gradescope mobile app. Student view of in-class file-upload activity.

Students were, overall, more successful in Zoe's course after she used Gradescope: Their average score increased and the DFW rate decreased (Table 2).

Table 2. Gradescope course outcomes by course: Business statistics.

Before or after Gradescope	Total no. students	Average DFW	Average score (Canvas)	Average Gradescope submission score
Before	177	8%	74	81
After	300	3%	91	87
After	1,078	4%	87	81

Note. DFW = Drop/fail/withdraw rate.

Additionally, Zoe created one rubric on Gradescope to provide consistency for grading: "practically the same rubric for both sides and for every activity." She also altered the intensity of the grading from having 20 parts to grade to simply grading each side of the page. Zoe changed the label of the term to protect against cheating from another term. Without the assistance from Gradescope, Zoe indicated that she would have needed to eliminate this interactive group work during class because of the challenges: "We probably could not have kept doing it…with turning in paper."

Faculty 2: Andrea

Andrea volunteered to use Gradescope in her Organic Chemistry classes as part of the university's spring 2023 pilot phase. The local teaching center contacted Andrea to test out this new technology with her classes because of her class sizes. In her interview, Andrea described facing four challenges to providing students with timely feedback: (a) large amount of grading time required with tight deadlines, (b) quality control among multiple graders, (c) transferring grades from paper to online within the LMS, and (d) subjective identification (by graders) of most-missed questions. After the pilot, Andrea continued using Gradescope in her spring 2024 Organic Chemistry class and added it to her fall 2023 Fundamentals of Chemistry and Biochemistry II course.

According to provided course materials, Andrea used Gradescope with her quizzes and exams that consisted of a combination of open essays, drawing chemical structures, and fill-in-the-blank questions with no multiple choice (see Figure 2). In her interview, Andrea discussed complications associated with the large variety of questions: "There are a lot of question types and a lot of moving parts…[which] requires 20 graders." Prior to using Gradescope, Andrea created a general rubric where each feedback comment was assigned an alphanumeric code for reference by her grading team. Each grader was given the rubric, the feedback key, and one exam question to grade for the hundreds of students.

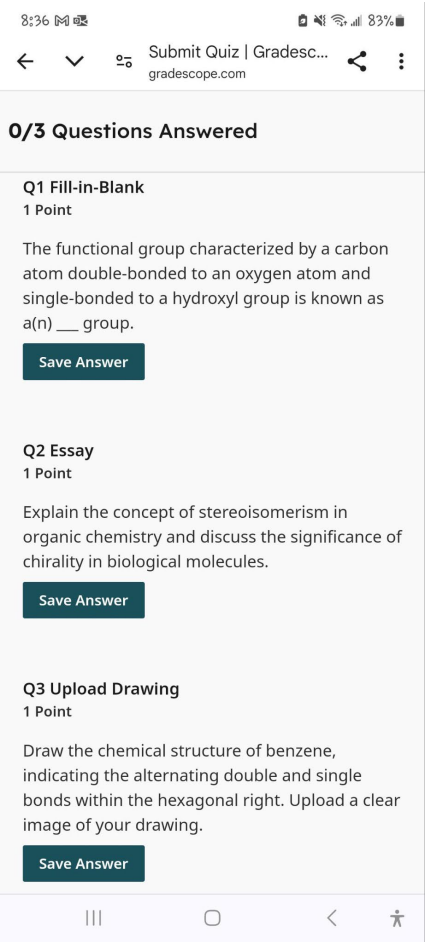

Figure 2. Gradescope mobile app quiz. Student view of fill-in-the-blank, short essay, and file upload questions.

When giving feedback, the grader recorded the appropriate code instead of writing out the same response multiple times. Students were then given the key along with their exam or quiz. In her interview, Andrea shared that quality control and grading consistency in the exams and quizzes were a challenge. Before using Gradescope, she gathered all the graders in a physical room for an 8-hr grading session. Graders unable to attend had to schedule individual grading time with the faculty member, which led to her spending more time on exam grading, thereby increasing time spent on logistics of the course. During those grading sessions, graders frequently asked her clarifying questions (e.g., "I have this weird answer—would you give him any points?"). Although these meetings helped with grading consistency and fidelity, graders often struggled with awarding partial credit.

Further, graders would share their perceptions of the most-missed question(s). During the lecture class following each quiz, Andrea used this information to address misunderstandings about the material. The transition actually helped Andrea increase class time spent on students' learning rather than on rudimentary instructional tasks. It took a large chunk of time to pass out the graded exams: "that's 20 minutes of lost instructional time"; whereas with Gradescope, "as soon as I post grades, they get to see their exams. We don't lose them if they're uploaded, [and] students can look at them anytime."

It is clear from Andrea's interview that Gradescope addressed the four challenges to providing timely feedback to her students. First, Gradescope permitted her to reduce the 8-hr in-person meetings to about half the time for the graders. The instructor's time was approximately the same, but it shifted from grading to preparation using Gradescope. This shift is a positive benefit as the instructor still had the same time input, but the 20 graders had a dramatically reduced workload. Additionally, there was the added benefit of graders being able to "grade from anywhere."

Second, Andrea believed that students' grades were higher and better represented student learning: "more accurate reflection of what students have been putting down on the page." Gradescope made it easier to give partial points, likely owing to a checkbox: "no amount of point taking off was too small because all it was, was clicking a box and they didn't physically have to write it out." Because Gradescope provided a mechanism for creating and applying a more detailed rubric, this made grading more consistent and seamless for graders. If the instructor did not agree with the level of points being deducted, it was simple to change the point deduction in the rubric to regrade all the problems that were graded perhaps too harshly or too easily.

Third, because Gradescope was integrated into the LMS, it made grade reporting easier as "it was automatic." The students did not need to wait until the next week in discussion to physically pick up their exams but instead could review their grades as soon as they were posted online. Fourth, Andrea appreciated the ability to analyze student learning across questions through the statistics features in Gradescope analytics: Not only did it allow her to "pull out the common mistakes" but "sometimes it would actually shock me that the most common wrong answer wasn't what I expected." Andrea described that having insight into the analytics behind that wrong answer provided her with accurate information to address misconceptions in the class.

According to the course data, the DFW rate decreased in Andrea's Organic Chemistry courses (see Table 3). Spring 2023 was Andrea's first time teaching the Fundamentals of Chemistry and Biochemistry II course, but students' performance aligned with the general outcomes for students within that course. Andrea reported that overall, Gradescope created greater student satisfaction with the course as a result of the quick and effective feedback: "I think the grading is more fair and it increases their [students'] satisfaction because they're getting better feedback…exams get returned to them right away [and] can't get lost." "I've never had a student have any complaints," stated Andrea. "We have a lot less actual complaints about grading." Andrea shared, "The students like it [because]…they have access to the regrade" and can easily see whether assignments have been submitted or not (See Figures 3 and 4).

Table 3. Gradescope course outcomes by course: Chemistry.

Course	Before or after Gradescope	Total no. student	Average DFW	Average score (Canvas)	Average Gradescope submission score
Fundamentals of Chemistry and Biochemistry II	After	120	5%	92	80
Organic Chemistry I	Before	513	39%	72	74
Organic Chemistry I	After	1,023	24%	81	71

Note. DFW = Drop/fail/withdraw rate.

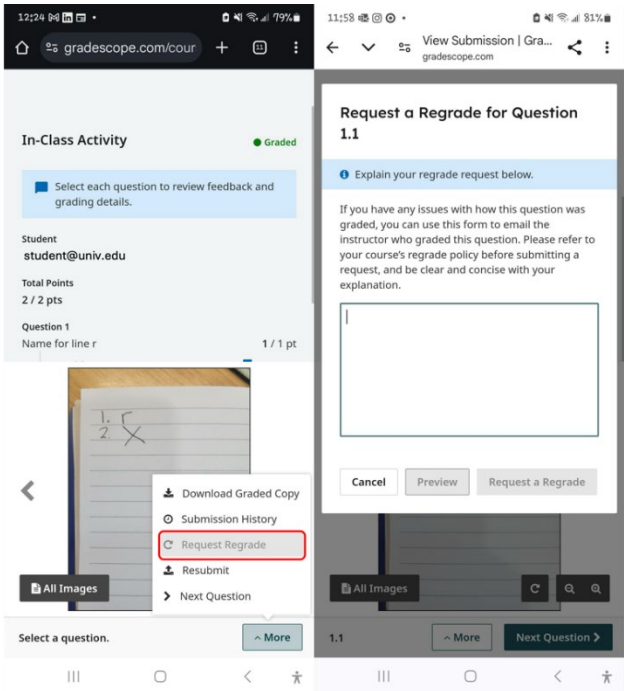

Figure 3. Gradescope mobile app. Student view of feedback and regrade request.

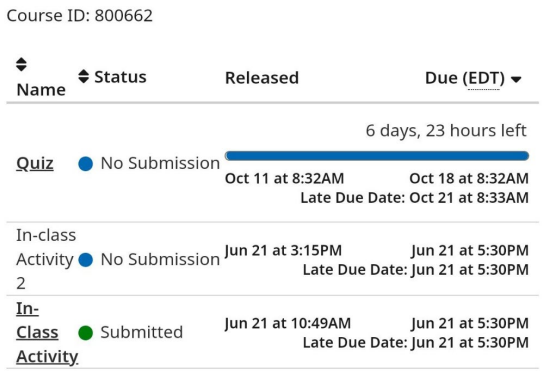

Figure 4. Gradescope mobile app. Student view of submitted and unsubmitted assignments.

Primary Themes

The two themes across both faculty were preparing to use the tool prior to launching it in the classroom and having a specific purpose for using the tool. Both cases highlighted the need for preparation time in learning how to use the tool and for creating rubrics to assist with the grading process. Andrea spent preparation time learning the tool, creating a rubric prior to using the tool with her classes, and preparing the assessments before being able to grade. In addition to making the rubric in Gradescope, Andrea's workload included cutting off the staples and uploading the exam or quiz pages into Gradescope for the students. Since Zoe used Gradescope for in-class activities, the focus of her preparation was on learning how to use the tool in class and building the in-class activities with rubrics.

Both faculty had specific purposes for Gradescope. Andrea used the tool to streamline the grading process of exams and quizzes and Zoe used the technology to continue providing students with in-class group work. In both cases, a benefit was the timely relay of feedback to their students.

Gradescope allowed Andrea's graders to digitally grade from anywhere and reduced the time spent grading. The tool's analytic feature additionally provided Andrea with data on the most-missed questions, which enhanced her ability to adjust her teaching to deepen her students' understanding on topics they found confusing. Zoe's specific purpose for using Gradescope was to make grading her paper-based activities sustainable. Gradescope provided her with the right tool to give students timely feedback digitally without the difficulties of handling paper-based assignments.

Discussion

Findings from this multicase study provide guidance on using this educational tool within a large-classroom setting:

1. Start early.
2. Build in time to learn the tool, prepare the materials, and incorporate the tool into the class design.
3. Create detailed rubrics for grading consistency.
4. Use the tool to meet a specific need within the class.

Gradescope provides a sustainable and timely student feedback process for in-class paper-based activities. The simplicity of the tool allows for refined grading, such as giving partial points. Grading can be done anywhere and any time since it is electronic. Grader access can be limited to the activities within the tool. The students can see the feedback as soon as it is posted. The technology eliminates wasted class time from handing back paper assignments while also providing data analytics on the most-missed questions.

Faculty face a wide range of challenges in large lecture courses, such as student depersonalization, reduced motivation, and difficulties in managing and grading assignments efficiently. To counteract these challenges, faculty can use feedback, which is a crucial element of learning for students who struggle with motivation and self-regulation (Jacobi, 2018; Lim & Kim, 2002-2003). However, timely and effective feedback can be challenging for faculty (Lim & Kim, 2002-2003. With the evolution of digital tools, faculty have new options to give learners quality feedback (Jensen et al., 2021) while managing the inherent issues of giving feedback in a large lecture classroom (Bernius et al., 2022).

This study explored the efficacy of Gradescope as one such digital tool for providing quality feedback in two large lecture classrooms. Two cases were presented of faculty who used Gradescope

to streamline their grading process, enhance their feedback quality, and improve their students' learning outcomes. Multiple cases demonstrate that Gradescope is an effective technology that has helped faculty overcome some of the specific challenges associated with large lecture courses, including streamlining the grading process, enhancing feedback quality, and maintaining grading consistency across large numbers of students and graders.

Looking at the course results, DFW rates decreased and average LMS grades increased during the semester in which the faculty used Gradescope. Faculty also reported increased grading efficiency and improved student satisfaction, attributing some of this to the use of Gradescope. However, the use of Gradescope was not entirely responsible for the positive impacts. Faculty also adapted their teaching methods and assessment strategies to integrate Gradescope effectively. This included changes in their feedback through the use of grading rubrics, the method of handling assignments, and adjustments in the instructional design to accommodate the tool's capabilities.

These two cases showcase the potential of using a tool such as Gradescope. These grading technology tools can help facilitate the provision of efficient and consistent feedback to students in large lecture classrooms. However, faculty should be guided in terms of pedagogical usage of the tool. Faculty should first clearly identify learning objectives and define what they are trying to achieve with the use of Gradescope or a similar grading tool (e.g., reducing grading time, improving feedback quality, or enhancing student engagement).

Since one of the key benefits of Gradescope is its ability to standardize grading through customized rubrics, faculty need to take the time to develop detailed, clear rubrics that can address the specific needs of their courses and assignments. This will ensure consistency in grading but will also help in providing targeted feedback to students. If using teaching assistants, faculty need to ensure they have been trained properly to apply the rubrics consistently and manage their sections of the grade book securely. Additionally, Gradescope simplifies adoption by providing a single point of interaction for faculty assessing student artifacts.

These activities should also be designed to encourage active participation. Gradescope can be used to foster more interactive and engaging learning environments by providing immediate and constructive feedback. Faculty might also consider how they could use Gradescope with a flipped classroom approach to increase engagement (Srinivansan et al., 2018). Faculty could use Gradescope for a scavenger hunt during orientation week or to upload photos of their chemistry lab experiment results.

These activities may also provide benefits to smaller scale classrooms. Mechanisms for providing clear feedback enable students to improve their learning within and beyond a specific activity and faculty to refine their assessment approach. Specifically, the use of clear criteria in practice facilitates clarification of the task at hand and the importance and contingency of evaluative criteria, and it introduces reflective opportunities to transform teaching and learning practices (Torrance, 2012). In this regard, Gradescope in smaller scale classrooms may provide opportunities for formative and transformative refinements to existing practices from the perspective of both faculty and students. Given these results, we suggest that other institutions consider similar technology, for training faculty and teaching assistants, and potential opportunities for further research on digital tools in education.

Despite the promising findings in the present study, there are limitations to its scope. The study used a multiple-case design that provides a detailed account of two faculty members' use of Gradescope in real-world classrooms, which is the purpose behind using this design. Additional cases would provide more insights into the many ways that faculty use Gradescope. A limitation is the combined deidentified student scores from Gradescope activities and deidentified final course grade data, making it impossible to make comparisons on individual students' progress or revisions they may have made. This was required by the IRB protocol. Since the goal was to focus on faculty members'

experience and the impact on deidentified student grades, the current study provides an initial baseline and inspiration for more in-depth analysis of the student experience.

Further research of these types of grading tools should examine how such tools can enhance active learning, provide timely and constructive feedback, and potentially increase course completion rates by making large classes feel smaller and more personalized. Additionally, comparing courses using Gradescope with other courses inside the same discipline that do not use it would be beneficial to learn more about the impact of Gradescope on course outcomes. Although the benefits of Gradescope are highlighted, it is also important to discuss the limitations or challenges that emerged during its implementation. This may include resistance from faculty or students, technical issues, or challenges in adapting existing courses to new digital formats.

References

Aljaloud, A., Gromik, N., Billingsley, W., & Kwan, P. (2015). Research trends in student response systems. *International Journal of Learning Technology, 10*(4), 313–325. https://doi.org/10.1504/IJLT.2015.074073

Bernius, J. P., Krusche, S., & Bruegge, B. (2022). Machine learning based feedback on textual student answers in large courses. *Computers & Education, 3*, 1–16. https://doi.org/10.1016/j.caeai.2022.100081

Betti, A., Biderbost, P., & Garcia Domonte, A. (2022). Developing students' "soft skills" through the flipped classroom: Evidence from an international studies class. *International Studies Perspectives, 23*, 1–24. https://doi.org/10.1093/isp/ekab014

Bolliger, D. U., & Des Armier, D. (2013). Active learning in the online environment: The integration of student-generated audio files. *Active Learning in Higher Education, 14*(3), 201–211. https://doi.org/10.1177/1469787413498032

Bouchard, P. (2009). Some factors to consider when designing semi-autonomous learning environments. *Journal of e-Learning*, *7*(2), 93–100.

Branca, S., & Slusser, E. (2022). Through a more discerning lens: Understanding college student expectations and experiences over the course of a semester. *College Student Journal, 56*(2), 190–196.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3,* 77–101. https://doi.org/10.1191/1478088706qp063oa

Cash, C. B., Letargo, J., Graether, S. P., & Jacobs, S. R. (2017). An analysis of the perceptions and resources of large university classes. *CBE – Life Sciences Education, 16*, 1–12.

Clayton, K., Blumberg, F., & Auld, D. P. (2010). The relationship between motivation, learning strategies and choice of environment whether traditional or including an online component. *British Journal of Educational Technology*, *41*(3), 349–364. doi:10.1111/j.1467-8535.2009.00993.x

Gok, T. (2011). An evaluation of student response systems from the viewpoint of instructors and students. *The Turkish Online Journal of Educational Technology, 10*(4), 67–83.

Habel, C., & Stubbs, M. (2014) Mobile phone voting for participation and engagement in a large compulsory law course. *Research in Learning Technology, 22*, 1–15. http://dx.doi.org/10.3402/rlt.v22.19537

Hartnett, M. (2012). Relationships between online motivation, participation, and achievement: More complex than you might think. *Journal of Open, Flexible, and Distance Learning*, *16*(1), 28–41.

Higgins, R., Hartley, P., & Skelton, A. (2002). The conscientious consumer: Reconsidering the role of assessment feedback in student learning. *Studies in Higher Education, 27*(1), 53–64. https://doi.org/10.1080/0307507012009936 8

Hollister, B., Nair, P., Hill-Lindsay, S., & Chukosie, L. (2022). Engagement in online learning: Student attitudes and behaviors during COVID-19. *Frontiers in Education, 7*, 1–16. https://doi.org/10.3389/feduc.2022.851019

Jacobi, L. (2018). What motivates students in the online communication classroom? An exploration of self-determination theory. *Journal of Educators Online*, *15*(2), 1–16.

Jensen, L. X., Bearman, M., & Boud, D. (2021). Understanding feedback in online learning—A critical review and metaphor analysis. *Computers & Education, 173*, 1–12. https://doi.org/10.1016/j.compedu.2021.104271

Johnson, B. E. (2011). The speed and accuracy of voice recognition software-assisted transcription versus the listen-and-type method: A research note. *Qualitative Research, 11*(1), 91–97. https://doi.org/10.1177/1468794110385966

Keller, J. M., & Suzuki, K. (2004). Learner motivation and e-learning design: A multinationally validated process. *Journal of Educational Media, 29*(3), 229–239.

Kokkelenberg, E. C., Dillon, M., & Christy, S. M. (2008). The effects of class on student grades at a public university. *Economics of Education Review, 27*(2), 221–233. https://doi.org/10.1016/j.econedurev.2006.09.011

Kuh, G. D., Kinzie, J., Buckley, J., Bridges, B., & Hayek, J. C. (2007). *Piecing together the student success puzzle: Research, propositions, and recommendations. ASHE Higher Education Report, 32*(5). Jossey-Bass.

Kuh, G. D., Kinzie, J., Schuh, J. H., Whitt, E. J., & Associates. (2005). *Student success in college: Creating conditions that matter*. Jossey-Bass.

Lee, E., Pate, J. A., & Cozart, D. (2015). Autonomy support for online students. *Tech Trends, 59*(4), 54–61.

Lim, D. H., & Kim, H. (2002-2003). Motivation and learner characteristics affecting online learning and learning application. *Journal of Educational Technology Systems*, *31*(4), 423–439.

McKnelly, K. J., Howitz, W. J., Thane, T. A., & Link, R. D. (2023). Specifications grading at scale: Improved letter grades and grading-related interactions in a course with over 1,000 students. *Journal of Chemical Education, 100*(9), 3179–3193. https://doi.org/10.1021/acs.jchemed.2c00740

Merriam, S. B., & Tisdell, E. J. (2016). *Qualitative research: A guide to design and implementation* (4th ed.). Jossey-Bass.

Morosan, C., Dawson, M., & Whalen, E. A. (2017). Using active learning activities to increase student outcomes in an information technology course. *Journal of Hospitality & Tourism Education, 29*(4), 147–157. https://doi.org/10.1080/10963758.2017.1382369

Mulryan-Kyne, C. (2010). Teaching large classes at college and university level: Challenges and opportunities. *Teaching in Higher Education, 15*(2), 175–185. https://doi.org/10.1080/13562511003620001

Ng, M., & Newpher, T. M. (2021). Class size and student performance in a team-based learning course. *Journal of Undergraduate Neuroscience Education*, *20*(1), A49–A57.

Nja, C. O., Orim, R. E., Neji, H. A., Ukwetang, J. O., Uwe, U. E., & Ideba, M. A. (2022). Students' attitude and academic achievement in a flipped classroom. *Heliyon, 8*(2022), 1–14. https://doi.org/10.1016/j.heliyon.2022.e08792

Pascarella, E. T., & Terenzini, P. T. (2005). *How college affects students: A third decade of research* (Vol. 2). Jossey-Bass.

Paulus, T., Lester, J. N., & Dempster, P. G. (2014). *Digital tools for qualitative research*. Sage.

Politis, J., & Politis, D. (2016). The relationship between an online synchronous learning environment and knowledge acquisition skills and traits: The Blackboard Collaborate experience. *The Electronic Journal of e-Learning*, *14*(3), 196–222.

Srinivasan, S., Gibbons, R. E., Murphy, K. L., & Raker, J. (2018). Flipped classroom use in chemistry education: Results from a survey of postsecondary faculty members. *Chemistry Education Research and Practice*, *19*(4), 1307–1318.

Stoerger, S., & Kreiger, D. (2016). Transforming a large-lecture course into an active, engaging, and collaborative learning environment. *Education for Information, 32*(1), 11–26.

Torrance, H. (2012). Formative assessment at the crossroads: Conformative, deformative, and transformative assessment. *Oxford Education Review, 38*(3), 323–342 https://doi.org/10.1080/03054985.2012.689693.

Tuma, F. (2021). The use of educational technology for interactive teaching lectures. *Annuls of Medicine and Surgery, 62*(2021), 231–235. https://doi.org/10.1016/j.amsu.2021.01.051

Turnitin. (2024a). *Can multiple people grade at once.* https://guides.gradescope.com/hc/en-us/articles/21586387201421-Can-multiple-people-grade-at-once#:~:text=Yes!,among%20the%20graders%20for%20you

Turnitin. (2024b). *Gradescope*. https://www.gradescope.com/

Visser, L., Plomp, T., Amirault, R., & Kuiper, W. (2002). Motivating students at a distance: The case of an international audience. *Educational Technology, Research and Development, 50*(1), 94–110.

Yin, R. K. (2016). *Qualitative research from start to finish* (2nd ed.). Guilford Press.

Yin, R. K. (2018). *Case study research and applications: Design and methods* (6th ed.). Sage Publications, Inc.