



# Technology Evaluation Report: Gradescope

## Report Contents

<a href="#">Executive Summary</a>	..... 2
<a href="#">Goal</a>	..... 5
<a href="#">Methodology</a>	..... 6
<a href="#">Differences from Crowdmark evaluation</a>	..... 8
<a href="#">Findings</a>	..... 8
<a href="#">Instructor motivations</a>	..... 8
<a href="#">Instructor and TA response to using Gradescope</a>	..... 9
<a href="#">Student response to using Gradescope</a>	..... 15
<a href="#">Recommendations</a>	..... 18
<a href="#">1) Carefully consider creation of rubrics</a>	..... 18
<a href="#">2) Highlight time savers for new TAs</a>	..... 19
<a href="#">3) Have students use the application directly</a>	..... 20
<a href="#">4) Raise the bar for online student regrade requests</a>	..... 20
<a href="#">5) Caution TAs to actively stay mindful when grading</a>	..... 21
<a href="#">6) Build in a fail-safe when multiple graders work in parallel</a>	..... 21
<a href="#">7) Include and use the artificial intelligence features</a>	..... 21
<a href="#">8) Extract lessons from the Computer Science workflow</a>	..... 22
<a href="#">Appendices</a>	..... 23
<a href="#">Appendix A - Instruments</a>	..... 23
<a href="#">Appendix B - Compiled Student Data</a>	..... 25

**Note:** This report summarizes outcomes of a specific UBC pilot. Findings do not reflect broader or official UBC opinions about the learning technology evaluated.

---

## Executive Summary

This evaluation sought feedback from UBC instructors, teaching assistants (TAs), and students piloting Gradescope, an application for online grading.

Once student submissions are uploaded, Gradescope allows for distributed grading by instructional teams, meaning instructors and TAs can view, assign rubrics, and leave comments on different questions or aspects of the same submission simultaneously. The rubrics used for grading may be created in advance, but Gradescope also records all grader rubrics and comments on-the-fly. Gradescope offers artificial intelligence features for grading as well, which can automatically group similar responses for three question types (multiple choice, formula, and fill-in-the-blank). Gradescope can distribute graded submissions directly back to students online (if they register accounts), or instructional teams can download the graded submissions for external digital handback.

Five instructors, 16 TAs, and 39 students provided feedback on Gradescope during the consultation period that ran between January and May 2018. The seven courses in the official pilot were from Applied Science and Science. Two were first-year courses, two were second-year courses, and three were third-year courses. All courses had larger class sizes, with registration ranging from 120-400 students and TA groups of 6-20 per course.

This evaluation differed from the earlier evaluation of Crowdmark—another application for online grading piloted at UBC—in three notable ways:

1. TA/student feedback was included
2. Use was limited mostly to the Computer Science department
3. Issues raised in the Crowdmark evaluation regarding the digitizing workflow were largely absent, since Computer Science already had a robust process in place

Instructors and TAs responded positively to Gradescope. Four out of five instructors felt Gradescope saved them time over paper grading, with one instructor saying it took about the same amount of time. All 16 TAs reported they would choose Gradescope for future grading over other online grading applications (though few had such experience) or working on paper. Both groups also rated learnability highly.

---

Instructional teams reported the application increased efficiency by offering rubric hotkeys, removing the overhead of paper shuffling, and adding flexibility in when and where grading could occur. TAs appreciated the anonymity, not only in obscuring student names from them but obscuring their names from students. Most instructors and TAs commented on improved consistency and confidence in grading with Gradescope, largely due to the ease of editing rubric descriptions and weights on-the-fly. The detailed nature of the rubrics contributed to a perceived higher quality of student feedback (since students see these rubrics as part of handback) as did the comments TAs could enter and reuse.

Gradescope offered a couple of extra features instructors weren't expecting but also found beneficial. All instructors tried and were impressed by the usefulness and general accuracy of the artificial intelligence features, and all instructors and several TAs noted the usefulness of tracking their grading progress with the Gradescope analytics.

As for drawbacks, many instructors found at times it was challenging to set up sensible rubrics with Gradescope's limitations, particularly since rubric items could only be binary and yet too many items created a less efficient and potentially more confusing grading experience. A recurring issue mentioned by both instructors and TAs was occasional overwriting of a grade, meaning work was duplicated with two TAs spending time grading the same answer. A few other minor application quirks also came up with both groups of users, notably the inability to apply a consistent rubric filter in reviewing graded exams and some inconveniences in how students review graded submissions.

Instructors and TAs also acknowledged potential workflow issues with using Gradescope outside Computer Science, which has worked out a relatively smooth departmental process for scanning, anonymizing files for upload, and mapping downloaded files for distribution. One instructor-specific concern with Gradescope related to the rubric hotkeys being too easy to use, something echoed by a few TAs who said keeping focused was important TA advice. Finally, one TA-specific pain point was that students who made regrade requests in Gradescope did not always take the care they normally might in requesting in person.

Students responded well to Gradescope, with all survey respondents choosing a positive rating of their overall application experience. The majority of students additionally reported

---

faster turnaround times and better feedback received for their Gradescope work. Students commented on Gradescope providing them increased clarity in how their work was graded, more convenient access to their uploaded work, grades, and feedback, and easier regrade requests. The only significant downside mentioned was that a couple students struggled to understand rubrics for more complex questions with multiple ways to gain or lose marks.

Based on these outcomes, recommendations for how Gradescope could best be implemented at UBC include:

1. Carefully considering the creation of rubrics in the application to maximize efficiency for graders and minimize confusion for students
2. Highlighting time savers for TAs new to Gradescope so they are aware of the shortcuts available to them from the start
3. Having students use the application directly to have access to all the application features (e.g., online regrade requests)
4. Raising the bar for online student regrade requests (if using the application directly), forcing students to clearly articulate why their assessment deserves another look
5. Cautioning TAs to actively stay mindful when grading to avoid mindlessly clicking hotkeys that don't apply
6. Building in a fail-safe when multiple graders work in parallel to avoid accidental overwriting of grades (unless the application builds in better prevention)
7. Including and using the artificial intelligence features to save time on questions that can utilize it
8. Extracting lessons from the Computer Science workflow (for other departments) in effectively printing, scanning, uploading, and handing back assessments

Implementing Gradescope with these recommendations may help resolve some of the concerns brought forward by instructional teams and students and improve future users' perceptions of the tool's pedagogical value.

---

## Goal

This evaluation sought feedback from UBC instructors, teaching assistants (TAs), and students piloting Gradescope, an application for online grading.

Once student submissions are uploaded (using original digital files or scanned files of paper-based work), Gradescope allows for distributed online grading by instructional teams, meaning instructors and TAs can view, assign rubrics to, and leave comments on different questions or aspects of the same submission simultaneously. For question-based submissions, responses to each question are grouped together, so answers for the whole class can be graded in sequence.

The rubrics used for grading may be created in advance, but Gradescope also records all grader rubrics and comments on-the-fly. This means, as grading proceeds, updated rubrics and comments are available for the entire instructional team to use. Gradescope also offers artificial intelligence (AI) features, which can automatically group similar responses for three question types (multiple choice, formula, and fill-in-the-blank). This allows graders to verify and apply rubrics to groups of answers, rather than grade each individually (though individual grading of answers is still an option).

Gradescope can distribute graded submissions directly back to students online (if they register accounts), or instructional teams can download the graded submissions for external digital handback. Downloaded submissions include a cover sheet with the grading outcomes, which students can use to reference the individual parts of their submission (e.g., in reviewing question #3, the student can find the rubrics assigned to that question on the cover sheet, then consult question #3 in the exam).

This report will detail the methodology used in evaluating Gradescope (including how this differed from the earlier evaluation of Crowdmark, another online grading application piloted at UBC<sup>1</sup>), why instructors chose to use it, the potential pedagogical value identified by people in the pilot, and how the tool could best be implemented at UBC.

---

<sup>1</sup> The Crowdmark official pilot ran in the 2016/17 W2 term, and the evaluation report was completed in July 2017.

---

## Methodology

The Gradescope official pilot ran during the 2017/18 W2 term in seven courses, although many instructors had also used Gradescope in prior terms. Five instructors were interviewed at or near the end of the term. Each interview consisted of 11 questions (see [Appendix A.1](#)), directly or indirectly addressed during in-person meetings.

TAs in all seven courses were invited to complete an online survey or participate in one of three in-person focus groups near the end of term; 16 provided feedback, with 14 of these doing so through a focus group. At the time of feedback, all TAs had used Gradescope for a minimum of one exam, though several had more experience. The focus groups and survey covered 8 questions (see [Appendix A.2](#)), with quantitative responses always submitted privately<sup>2</sup> and qualitative responses given privately by survey respondents but forming the basis for discussion in focus groups.

Students who set up their own Gradescope accounts (done in Course 7 only) were asked to respond to an online survey near the end of term; 39 responded prior to May 2018. At the time of the survey, all students had used Gradescope multiple times to submit assignments and receive grades and instructional team feedback for assignments, exams, and quizzes. The student survey included 6 questions in total (see [Appendix A.3](#)).

The types of courses in the official pilot came from the Faculty of Applied Science and Faculty of Science (Computer Science aka CS). Two were first-year courses, two were second-year courses, and three were third-year courses. All courses involved larger class sizes, with registration ranging from 120-400 students and TA groups of 6-20 per course. Summarized use cases for each course are provided below; the only significant difference in application use was that one course used student accounts directly in Gradescope<sup>3</sup>.

---

<sup>2</sup> TAs in focus groups filled out a 5-question survey individually and anonymously at the start of the session, and online survey respondents replied anonymously to these 5 questions.

<sup>3</sup> Even with Gradescope's U.S.-based servers, this was compliant with privacy regulations because the students used unique, random identifiers assigned by the Computer Science department that could not be traced to any student identifying information. Gradescope servers never stored anything other than the students' anonymized submissions along with this identifier.

Faculty / Year	Students / TAs & Responses	Grading Context
<b>Course 1</b> Applied Science 1st Year	~300 students ( <i>not surveyed</i> ) _____ ~20 TAs 3 TAs responded	<ul style="list-style-type: none"> <li>Exams: 2 midterms, 1 final</li> <li>Students did not access Gradescope directly (received graded PDFs externally)</li> </ul>
<b>Course 2</b> Science (Computer Science) 1st Year	~250 students ( <i>not surveyed</i> ) _____ 14 TAs 0 TAs responded	<ul style="list-style-type: none"> <li>Exams: 2 midterms, 1 final</li> <li>Students did not access Gradescope directly (received graded PDFs externally)</li> </ul>
<b>Course 3</b> Science (Computer Science) 2nd Year	~400 students ( <i>not surveyed</i> ) _____ ~20 TAs 0 TAs responded	<ul style="list-style-type: none"> <li>Exams: 1 midterm, 1 final</li> <li>Students did not access Gradescope directly (received graded PDFs externally)</li> </ul>
<b>Course 4</b> Science (Computer Science) 2nd Year	~120 students ( <i>not surveyed</i> ) _____ 6 TAs 0 TAs responded	<ul style="list-style-type: none"> <li>Exams: 2 midterms, 1 final</li> <li>Students did not access Gradescope directly (received graded PDFs externally)</li> </ul>
<b>Course 5</b> Science (Computer Science) 3rd Year	~320 students ( <i>not surveyed</i> ) _____ 10 TAs 9 TAs responded	<ul style="list-style-type: none"> <li>Exams: 2 midterms, 1 final</li> <li>Students did not access Gradescope directly (received graded PDFs externally)</li> </ul>
<b>Course 6</b> Science (Computer Science) 3rd Year	~170 students ( <i>not surveyed</i> ) _____ 7 TAs 1 TA responded	<ul style="list-style-type: none"> <li>Exams: 1 final</li> <li>Quizzes: 4 quizzes</li> <li>Students did not access Gradescope directly (received graded PDFs externally)</li> </ul>
<b>Course 7</b> Science (Computer Science) 3rd Year	~357 students 39 students responded _____ 11 TAs 3 TAs responded	<ul style="list-style-type: none"> <li>Assignments: 5 assignments</li> <li>Exams: 2 midterms, 1 final</li> <li>Quizzes: 5 quizzes</li> <li>Students accessed Gradescope directly</li> <li>Students submitted assignments and received graded submissions in Gradescope</li> </ul>

---

## Differences from Crowdmark evaluation

This evaluation differed from the earlier evaluation of Crowdmark in three notable ways:

1. **Inclusion of student/TA feedback:** Due to the start of the Crowdmark evaluation coinciding with the end of term, direct TA/student feedback about Crowdmark was not available. With earlier notice for the Gradescope evaluation, planning for and collecting TA/student feedback before the end of term was possible.
2. **Single departmental context:** At the time of the Crowdmark evaluation, more instructors had piloted Crowdmark at UBC, allowing the application to be discussed in a wider range of departmental contexts. Gradescope use has been mostly limited to Computer Science, so instructor response was limited to that department.
3. **Digitizing workflow already established:** Issues raised in the Crowdmark evaluation regarding the paper workflow process (i.e., those around printing, scanning, uploading/anonymizing, and mapping to students) were largely absent in Gradescope discussions. This was due to the Computer Science department already having a robust process in place for digitization and digital handback of exams.

## Findings

### Instructor motivations

In moving to online grading, instructors stated one or more of the following motivations.

#### Better efficiency in grading

With paper-based student work, *"passing around paper and keeping track of which questions have been graded on which exams is an enormous pain"* for many instructional teams, particularly in large courses. Going online meant automating these processes, making it *"much easier to have a team of markers work in parallel"*.



---

## Guaranteed anonymity in grading

Some instructors *"don't want [TAs] to see the students' names before they grade"*, since *"research shows you extract all kinds of meaning from the name, no matter whether you know the person or not"*. Online grading could more easily enable anonymizing student work, so *"it's just a number on the exam"* graders see.

## Improved consistency in grading

Often in grading, instructional teams *"grade a bunch and then see that [an issue is] affecting more or fewer students than you expected"*. But it can be challenging and time-consuming to change the rubrics after-the-fact. Online grading could make it *"easy to spot grading errors/problems"* and then *"easily edit that rubric"* for the whole class.

## Higher quality of feedback for students

Some instructors wanted to communicate outcomes in detail with each student (even in large courses), so they could *"see exactly why they lost marks"*. Quality also included faster turnaround time *"since there's a lot of evidence that more rapid feedback gets more attention from students and so leads to better learning"*. Online grading offered potential for better feedback in a shorter timeframe.

## Security of digital archive

Online grading also means *"there's no concern about losing things"* in the grading process. As soon as student work is uploaded, it's in a secure location. Additionally, instructors could *"download marked exams for archival [purposes] after marking"* and keep these local copies as backups.

## Instructor and TA response to using Gradescope

Instructors and TAs responded positively to Gradescope. Four out of five instructors felt Gradescope saved them time over paper grading, with one instructor saying it took about the same amount of time. All 16 TAs reported they would choose Gradescope for future grading over other online grading applications (though few had such experience) or

---

working on paper. Both groups also rated learnability near the 'very easy to use' end of the 5-point scale (a '1' or '2'), with no instructor or TA finding it challenging.

### Increased grading efficiency

All instructors and TAs reported that Gradescope increased grading efficiency in a number of ways. First, *"the rubrics were linked to [hot]keys"*, meaning that graders could use a number on the keyboard to toggle an item on or off— *"we were all just sitting there with our fingers on the keys"*. Not dealing with paper translated to several smaller time savings in that TAs *"don't have to wait for that file somebody else is currently holding hostage"* to grade, *"there's no flipping to your particular page"* to get started each time, and *"it seems silly but there's no paper cuts"* to attend to. For some instructors, *"entering grades is another big win"*, since Gradescope automated this previously manual process (in which grades had to be recorded from paper exams to digital spreadsheets). And TAs liked the increased flexibility in when and where they could grade (when not grading with in-person groups), making it easier to *"do it in my downtime as opposed to having to set aside time"* to do it all at once.

### Higher comfort levels with anonymous grading

Most TAs appreciated the anonymity Gradescope provided, not only in obscuring student names from them but obscuring their names from students. *"As an undergraduate myself, I know people in the class, so I found that having it be anonymous....I never had to worry if I had to mark a friend"*. This anonymity applied to regrade requests in the application<sup>4</sup> too—*"you don't see which TA is responding"*—which was perhaps even more important. *"Students sometimes were incensed"* by the grade received for a question, and TAs *"don't want the personal tensions"* they may have encountered if students knew which of the TAs they corresponded with or that it was a TA at all (and not an instructor).

### Improved grading confidence and consistency

Most instructors and TAs commented on improved consistency in grading with Gradescope. Typically, the instructor or a designated TA set up rubrics ahead of time, but as grading progressed, graders would edit these rubrics, e.g., *"I would explain further in the*

---

<sup>4</sup> Note that only Course 7 had this option, as only these students had accounts in Gradescope.

---

*[rubric] items just in brackets saying how I applied it*" based on actual submissions. The ability to customize the rubrics in this way (and always interacting with the latest version to grade each submission) helped keep multiple graders on the same page. But also individually *"it gives you some degree of memory"* so *"it's much easier to keep 'this is what I'm doing' in mind"* as grading progressed, especially if sessions were spaced apart (as one TA noted, *"even reminding yourself what you did last weekend is very helpful"*).

While editing the rubrics helped specify their exact application, sometimes it also became clear *"for this particular question, people are not doing really well"* and editing of the rubric weighting was more appropriate. Gradescope *"allows you to do those sort of retroactive adjustments really really easily"*, so after an adjustment, *"I can regrade 300 exams with three keystrokes"*, a process on paper *"that is a huge amount of work"* (and for some instructors not worth the effort). All this on-the-fly massaging of rubrics led to increased *"confidence in the quality of our marking"* on behalf of some instructors and TAs that *"the grades I was giving were fair and a reflection of the student's work"* in the context of the class as a whole.

### **More detailed feedback for students**

For many instructors and TAs, the detailed nature of the rubrics also contributed to a perceived higher quality of feedback they were able to provide to students. Since the rubrics used are included as part of the handback, students see these, making it *"more explicit about what the points are"* for each question or section; on paper, *"we wouldn't write everything for every single question, not for [100+] students"*. Additionally, most TAs felt in Gradescope *"the feedback that you're able to give is a lot more rich"* to each individual because *"Gradescope keeps a record of all the comments that you leave"* (any of which can be reused) and *"if I'm typing, it only takes me a minute to write quite a lot of stuff"*. Between the rubrics and comments, *"the students do I think tend to get better feedback"* and often faster, e.g., *"we can give them back the quizzes in less than 48 hours"*.

### **Helpful artificial intelligence and analytic features**

All instructors tried and were impressed by the usefulness and general accuracy of the artificial intelligence features, both in cases of multiple-choice style and handwritten answers. While *"you do need to keep an eye on it"* and verify the automatic answer groupings

---

the AI created, *"the grouping saved a lot of work"* since rubrics could be assigned to entire groups at once. In one course, for a single question *"we marked 400 exams in a matter of minutes"*; another instructor noted *"that was just ridiculous"* in terms of time savings. For TAs who had first-hand experience double-checking the groups, several agreed that *"I would say that 95% of the time it was accurate"*, and *"it was easy to put [misplaced answers] into the right groups"*, then *"you just apply the rubric to all of them"*.

All instructors and several TAs also noted the usefulness of tracking their grading progress with the Gradescope analytics. On the instructor side, Gradescope provided insight *"to see how many exams each TA had marked"* and *"how much progress has been made on each problem"*, and *"that's really incredibly handy for making sure that I can see what's going on"* and intervening occasionally to nudge slower TAs or redistribute the workload. On the TA side, some thought knowing instructors *"have more insight into your grading"* and *"see even which TA graded"* which area provided extra assurance any errors they made would be caught before adversely affecting students. One TA also liked comparing their own progress to other TAs and felt seeing this *"gives you motivation to finish"* because *"it feels like a game"*.

### Limitations on setting up rubrics

As for drawbacks, many instructors found *"it's hard to enter a sensible rubric"* sometimes with the limitations Gradescope sets, particularly for more complex questions. Grading for each question had to be positive or negative, and rubric items were a) binary choices, b) listed off to the side in a single column when grading, and c) tied to number hotkeys. Once any question hit 10+ items, *"it means it's not going to fit on the screen at the same time, and you're not going to be able to use the keyboard to enter"* (thus negating the time savings of these shortcuts). And for problems with multiple ways of grading, *"the grading scheme got a little bit out of hand"*, at times confusing TAs, e.g., *"there was one where the profs had to go 'choose one of the following three options'"* and choosing more than one awarded extra points accidentally. In some instances, this encouraged instructors to *"lean toward marking schemes because of what is convenient to mark, instead of the right way to mark things"*.

---

## Accidental overwriting of grades

Another recurring issue mentioned by many instructors and TAs was the occasional overwriting of the same grade. While Gradescope does have a way to advance to the next ungraded exam, *“in some cases it would still go to the same one that someone else had opened already”*. It was unclear whether this was due to a race condition or an internal timeout related to the start of grading (i.e., clicking the first rubric). Either way, *“it won’t even tell you that someone else is grading that submission when you’re grading it”*, meaning work was occasionally duplicated, with two TAs spending time grading the same answer. Each course ended up implementing a workaround to minimize the chances of this occurring<sup>5</sup>.

## Minor application quirks

A few other minor application quirks also came up in conversations with instructors and TAs. One that multiple people mentioned related to filtering submissions by those with a particular rubric applied. The filter did not stick when using the next button—*“there is no way to filter and keep the filter persistent”*—meaning it was slightly more difficult to review the submissions rapidly. In a course that used group submissions, Gradescope did not alert when a student was in either multiple groups or no groups. Some instructors and TAs also commented on the inconvenient format of the graded submissions, specifically: 1) students with downloaded submissions had to navigate between a cover sheet with the rubrics and their actual exam answers—*“they have to go back and forth to see it well”*—and 2) students with online submissions saw no clear indicator of feedback comments<sup>6</sup>. Individual feature requests included the ability to manage the list of reusable comments, easily duplicate rubrics from one question to another (and between assignments/exams/quizzes), and see a timestamp of when grading was completed for each exam.

---

<sup>5</sup> See recommendations section for examples of these workarounds.

<sup>6</sup> Students only saw comments when they navigated directly to a question with comments; no overview was available at the time of the pilot, though Gradescope support said they are considering adding a higher-level indicator in a future release.

---

## Workflow issues with scanning

An issue brought up not specific to Gradescope was a problem with missed answers on scanned exams. Sometimes students would write lightly or not erase well enough, making it hard to decipher responses on the digital version. Additionally, *“if the student does put the answer somewhere else...on an electronic copy it’s a little longer sometimes to go through the pages and find out where”*. Infrequently, instructional teams would need to consult the paper copy itself, and *“that was a PAIN”* because *“these are exams that have had their staples chopped off”*, so sorting the huge pile of papers for a large course was time-consuming. Even if the answer was found in the digital version, *“that got a bit complicated”* switching between the screen where the rubric had to be entered and the actual answer location.

Both instructors and TAs acknowledged potential workflow issues to using Gradescope in other departments—issues the Computer Science department had already resolved at the time of the pilot. First of all, *“for people who don’t normally scan, that is going to be a slow-down without a doubt”*, whereas the CS department already had a process in place as well as quality scanners. Additionally, *“we’ve got scripts that can strip off the front cover”* (with the student identifying information) to expedite uploading anonymous submissions to Gradescope, and on the handback side, *“we match up the exam ID with the student ID through character recognition”* as well. Finally, for the course where students used Gradescope directly, these students relied on a CS-specific identifier for logging in that allowed them to circumvent any privacy issues, and *“I’m not sure if other faculties have the same thing”*.

## Higher chance of error with quicker grading

One instructor-specific concern with Gradescope related to the ease of the rubric hotkeys, that maybe *“it’s too easy to hit the wrong key”* or click one *“that won’t mean what [the TAs] think it means”* (especially since the rubric order can be rearranged on-the-fly with a simple drag and drop). A few TAs in fact brought the speed issue up in discussing advice for future TAs: *“sometimes I got so into it, and I was so fast with the keys”, “it becomes muscle memory”,* and *“if I see three correct submissions in a row, it’s tempting to just select ‘1-2-3-4-5’ for the next student”* without thinking. They agreed *“you have to make sure your mind’s still in it”*, though this was not seen as an issue from the TA perspective, just something to be careful about.

---

## Regrade requests sometimes overused by students

Finally, one TA-specific concern came up with regard to student regrade requests. In the course that allowed for these through Gradescope, a couple of TAs felt *“especially at the beginning of term, students tended to abuse this feature”*<sup>7</sup>. Removing the barrier of having to go directly to the instructor meant *“it probably seems like less of an effort”* and in some cases, *“that back and forth turned into a shouting match”* through the application. The instructor for this course did monitor the regrade requests and intervene when necessary, but some TAs still felt additional stress about having to directly (though very infrequently) receive and respond to *“somewhat disrespectful comments”* or attitude that they did not encounter with the paper-based request process<sup>8</sup>.

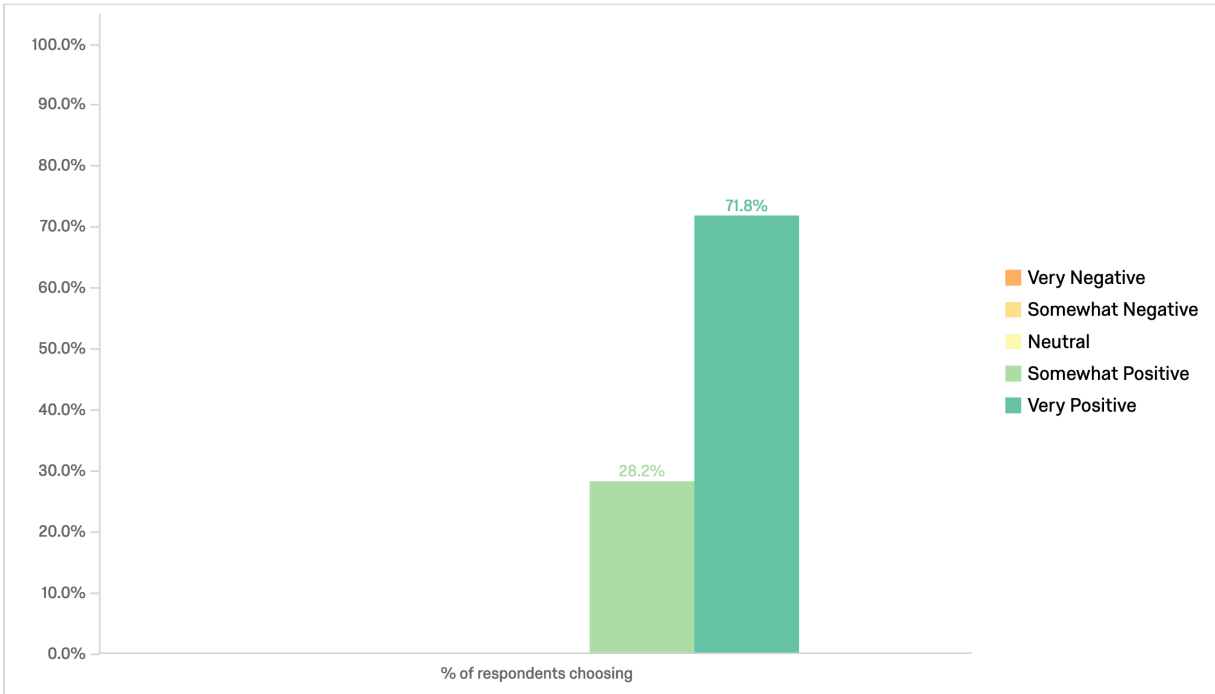
## Student response to using Gradescope

Though response rate was low in the one course that ran the survey, students who did respond responded well to Gradescope, with all choosing a positive rating for their overall application experience (N = 39).

---

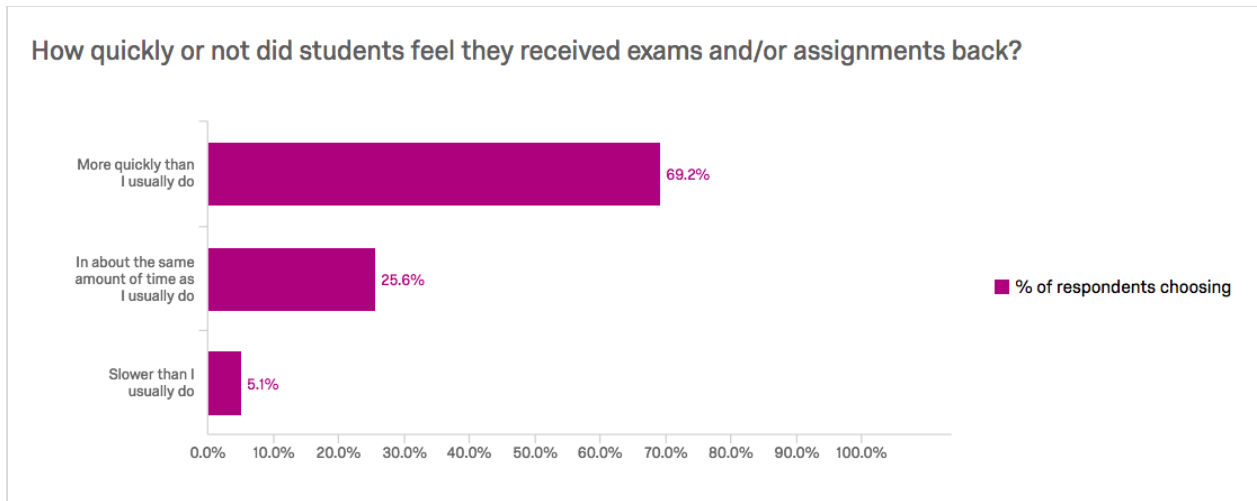
<sup>7</sup> Interestingly, those courses with regrade requests happening outside Gradescope generated split opinions on whether instructional teams received more or fewer regrade requests because of the detailed rubrics students saw.

<sup>8</sup> It should be noted TAs generally felt positive about the regrading experience, seeing this as a learning opportunity for students (and finding this easier to manage online than offline), and the majority of students were reasonable and polite in their requests.

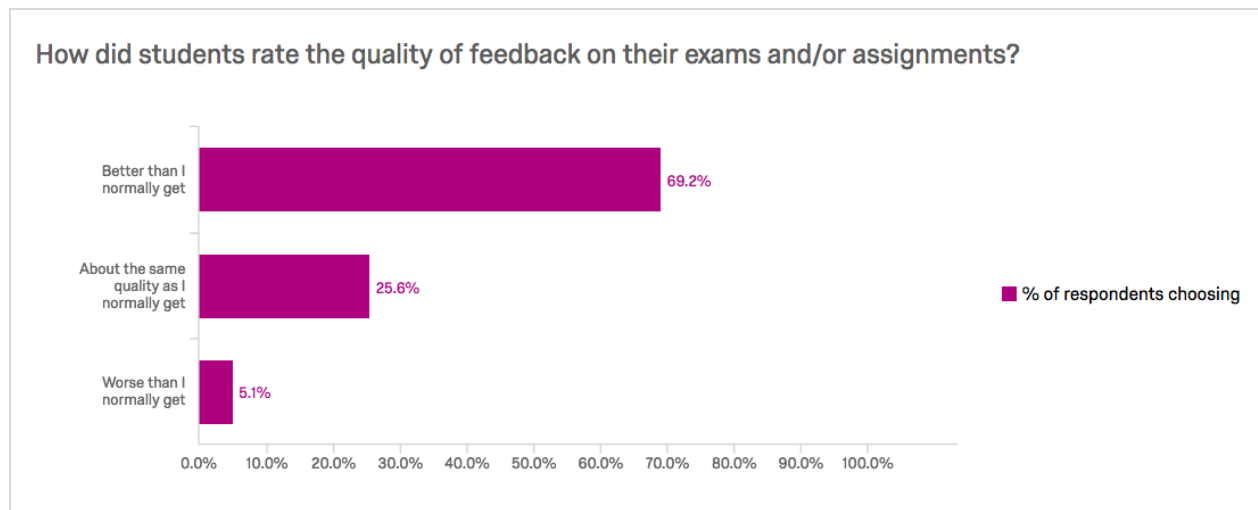


## Better turnaround time

The majority of students (69.2%) additionally reported faster turnaround times and better feedback received for their Gradescope work (N = 39).







## Clarity in grades received

By far, the top benefit for students who commented (see [Appendix B.1](#)) was finding greater clarity in why they had received the grades they did: *“I loved being able to go through my exam question by question and have a full comprehensive view of how each was marked”*. This seemed helped by the fact that *“the rubric breaks things down in a way that makes sense to the person viewing the exam”* in a side-by-side display, the feedback comments were typed (*“no issue ever with handwriting legibility”*), and the graders themselves had the willingness and ability to explain things well<sup>9</sup>. All this added up to a sense for many Course 7 students that Gradescope *“really helped me identify where my learning had gone wrong”* (or right).

## Convenience of submission and return

Another common point of appreciation for many students was the convenience offered by Gradescope. This applied to the ease of submitting assignments (*“we could submit at home”*) and accessing grades, feedback, and past submissions from anywhere (*“I can [access] this wherever I like with my laptop without having to carry a bunch of papers around”*). Additionally, *“getting a scanned copy of our exam + assignments + quizzes was awesome”*, allowing students a simple way to keep their own local digital archive.

<sup>9</sup> As one student astutely pointed out: *“While Gradescope is definitely beneficial, course organization and willingness to provide additional comments is the main contributor to good feedback and timely returns”*.

---

## Easier regrade requests

Finally, several students liked the streamlined nature of regrade requests in the application. *“Regrade requests were simple, completed in a timely manner, and with feedback, which was greatly beneficial.”*

## Confusion over complex rubrics

The only downsides of Gradescope students brought up (see [Appendix B.2](#)) were minor application quirks like seeing past courses and limitations on conversations during a regrade request. Two students also felt differently than many of their peers that the grading rubric was sometimes challenging to interpret, *“e.g., there would often be multiple options for the same answer that give different points (depending on the quality of the answer), but if none of those applied it was hard to see what exactly was wrong with your answer”*. This may have been an outcome of instructional teams trying to work around the rubric limitations noted earlier.

## Recommendations

Based on this pilot’s outcomes, these are some recommendations for how Gradescope could best be implemented at UBC to maximize its perceived benefits and minimize its perceived shortcomings.

1) Carefully consider creation of rubrics	
May address	<ul style="list-style-type: none"><li>• <a href="#">Increased grading efficiency</a> (instructor/TA benefit)</li><li>• <a href="#">Improved grading confidence and consistency</a> (instructor/TA benefit)</li><li>• <a href="#">Limitations on setting up rubrics</a> (instructor/TA issue)</li><li>• <a href="#">Clarity in grades received</a> (student benefit)</li><li>• <a href="#">Confusion over complex rubrics</a> (student issue)</li></ul>

Several instructors advised their peers to *“spend as much time ahead of time thinking about the rubric”* they plan to use in Gradescope. Rubrics need to take into account not just the best way to grade the work, but also Gradescope’s existing limitations, the grading efficiencies it offers, and the TA/student experience of interpreting the rubrics.

---

Some tips shared by instructional teams for this process included:

- When practical, **choose a positive or negative grading scheme based on what will be faster for TAs** to enter, e.g., *"I do negative when there is something I assume most students will do right or mostly right"* to save on clicks/keystrokes.
- Plan to **add an extra rubric item to indicate when an answer is entirely correct** (in a negative grading scheme) **or incorrect** (in a positive one), since in Gradescope *"to consider a submission marked, you need to assign some rubric item to it"*.
- Try to **avoid schemes that combine points in a single rubric** (i.e., one item wrong, two items wrong, three items wrong), as *"it's harder...to later on figure out who did what wrong in one particular item"*. Instead *"have individual rubric items for individual errors"* (or successes) to take advantage of the rubric-based filtering and analytics.
- Try to **avoid schemes that attempt to force TAs to choose among a group of rubric items** (e.g., choose one of the following three items), as this setup can be challenging for TAs to enter and students to interpret correctly.
- In the case of reaching **ten or more rubric items for a question, consider breaking the grading into two** parts, so half the items appear on the screen at a time.
- Set up an initial rubric description and weight for each item, but **modify descriptions and weights as grading progresses**. As new responses are encountered, *"edit that rubric item so it clearly fits under it"* (or doesn't) and for the weighting *"just adjust that in the end"*.

## 2) Highlight time savers for new TAs

<b>May address</b>	<ul style="list-style-type: none"><li>• <a href="#">Increased grading efficiency</a> (instructor/TA benefit)</li><li>• <a href="#">Higher quality of feedback for students</a> (instructor/TA benefit)</li><li>• <a href="#">Clarity in grades received</a> (student benefit)</li></ul>
--------------------	-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Instructors and TAs mostly agreed that not much training was required to get started grading in Gradescope. However, explicitly pointing out the ways in which TAs can

---

maximize their efficiency is important. TAs should be shown *“it’s even faster than just clicking, if you just type the numbers”* to select the rubrics, as well as use other hotkeys for navigating, since TAs said *“it’s not entirely obvious what every shortcut is”* in the beginning.

TAs stressed the importance of knowing *“you can reorder the rubrics”* and advised future TAs to use this to their advantage—*“you’d want the first ones to be the more frequent ones”* for any given question (though this needs to be coordinated among multiple graders). Additionally, as noted in the first recommendation, editing the rubrics to clarify how they have actually been applied can save TAs from having to keep these details in mind.

Highlighting the reusable comments is key here as well. TAs should try to write comments that are useful to a particular student problem but not so specific they cannot be reused for other students. And, if *“seeing how often we actually had to tell students that insight would be helpful”*, this would better be tracked as a rubric item instead of a comment. Gradescope does not allow for filtering or analytics based on comments (at least not as of this report).

### 3) Have students use the application directly

<b>May address</b>	<ul style="list-style-type: none"><li>• <a href="#">Convenience of submission and return</a> (student benefit)</li><li>• <a href="#">Easier regrade requests</a> (student benefit)</li></ul>
--------------------	----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Since the surveyed students used Gradescope directly and evaluated it as a positive experience, it follows that students may benefit from direct application use. This would allow them to submit their own work, view and download graded work at their leisure, and submit regrade requests online. However, this recommendation is dependent on having a privacy workaround (or permanent resolution regarding Gradescope’s server location).

### 4) Raise the bar for online student regrade requests

<b>May address</b>	<ul style="list-style-type: none"><li>• <a href="#">Regrade requests sometimes overused by students</a> (instructor/TA issue)</li></ul>
--------------------	-----------------------------------------------------------------------------------------------------------------------------------------

If Gradescope is used directly by students and regrading enabled, instructional teams might think about raising the bar for online regrade requests. For example, as is often the case for paper-based requests, telling students they have to provide written justification satisfying specific criteria as part of the online request (and be polite!) could force students

---

to stop and think before submitting requests. Additionally, TAs thought “*there probably should be someone overseeing it*” beside just the TA receiving the request. An instructor or lead TA should also be involved, in case discussion becomes more heated or the original grader has difficulty (re-)explaining the answer.

## 5) Caution TAs to actively stay mindful when grading

<b>May address</b>	<ul style="list-style-type: none"><li>• <a href="#">Improved grading confidence and consistency</a> (instructor/TA benefit)</li><li>• <a href="#">Higher chance of error with quicker grading</a> (instructor/TA issue)</li></ul>
--------------------	-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

As discussed earlier, TAs themselves acknowledged the need to actively “*stay focused*” when grading, since several found with the hotkeys that sometimes their “*fingers want to click the same thing every time*”. They touted this as good advice for future TAs using Gradescope.

## 6) Build in a fail-safe when multiple graders work in parallel

<b>May address</b>	<ul style="list-style-type: none"><li>• <a href="#">Accidental overwriting of grades</a> (instructor/TA issue)</li></ul>
--------------------	--------------------------------------------------------------------------------------------------------------------------

Each course ended up building in a fail-safe to deal with the issue of TAs overwriting one another’s grading. Unless Gradescope resolves the core problem, one of these approaches would be recommended. These included:

- Seating multiple TAs grading the same question next to each other (for in-person group grading) and visually or verbally checking in on one other’s progress
- Having two TAs grading the same question start at opposite ends of the exams and meet in the middle
- Using a “lock-in” rubric (a value-less rubric a TA immediately clicks after loading the page) to try to claim the grading

## 7) Include and use the artificial intelligence features

<b>May address</b>	<ul style="list-style-type: none"><li>• <a href="#">Increased grading efficiency</a> (instructor/TA benefit)</li><li>• <a href="#">Helpful artificial intelligence and analytic features</a> (instructor/TA benefit)</li></ul>
--------------------	--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

---

At this point, the AI features in Gradescope are enabled only at the second-tier level of pricing. These did prove useful to all instructors (though only one felt the absence of these features would significantly affect the usefulness of the overall application), so ideally adoption of Gradescope would include the option to use AI.

## 8) Extract lessons from the Computer Science workflow

<b>May address</b>	<ul style="list-style-type: none"><li>• <a href="#">Workflow issues with scanning</a> (instructor/TA issue)</li></ul>
--------------------	-----------------------------------------------------------------------------------------------------------------------



As noted, Computer Science has developed and refined a process for smoothly digitizing paper exams, including mostly automating the anonymizing and re-mapping of digital files to students. Working with CS to try to apply the lessons they have learned about this process in other departments could help ease implementation.

Implementing Gradescope with these recommendations may resolve some of the concerns brought forward by the instructional teams and students and improve future users' perceptions of the tool's pedagogical value.

---

# Appendices

## Appendix A - Instruments

### A.1) Instructor interview questions

1. What inspired you to use Gradescope?
2. Please explain the context in which you used Gradescope.
3. How did Gradescope help with your course?
4. What were the downsides or inconveniences of using Gradescope? (Also, if you did not choose to create student accounts, why not?)
5. What kind of training did you require or need to give graders? Students?
6. In terms of time commitment, how did grading with Gradescope (including printing, scanning, and training) compare to grading without Gradescope?
  - a. It took MORE time to grade with Gradescope
  - b. It took about the SAME amount of time either way
  - c. It took LESS time to grade with Gradescope
7. How would you rate your experience of learning to use Gradescope on a scale of 1-5, with 1 being very easy and 5 being very challenging.
8. When you needed help with using Gradescope, how did you get it?
9. On a scale of 0 to 10, with 10 being highest, how likely are you to recommend Gradescope to a colleague or friend for use in teaching and learning?
10. What advice would you give instructors considering using Gradescope?
11. How, if at all, can you see using Gradescope influencing the types of exam questions you choose in the future?
12. Anything else?

### A.2) TA focus group / survey questions

1. What level of experience do you have being a TA?
  - a. This is my first time being a TA
  - b. I've been a TA at least once before
  - c. I've been a TA multiple times before
2. What were you responsible for grading in Gradescope for this course (i.e., type of work you graded, # of exam questions or assignments, # of students)?

- 
3. How would you rate your experience of learning to use Gradescope on a scale of 1-5, with 1 being very easy and 5 being very challenging.
  4. For future grading, what would you prefer?
    - a. Gradescope online grading
    - b. Some type of online grading (but not with Gradescope)
    - c. Traditional paper grading
  5. On a scale of 0 to 10, with 10 being highest, how likely are you to recommend Gradescope for use in other courses you TA for?
  6. What did you like and/or not like about using Gradescope?
  7. What advice would you give TAs who may be using Gradescope in future terms?
  8. Any other feedback you'd like to provide on Gradescope or online grading generally?

### A.3) Student survey questions

1. I got my Gradescope exams and/or assignments back:
  - a. More quickly than I usually do
  - b. In about the same amount of time as I usually do
  - c. Slower than I usually do
2. The feedback I got on my Gradescope exams and/or assignments was:
  - a. Better than I normally get
  - b. About the same quality as I normally get
  - c. Worse than I normally get
3. Gradescope marking may be done anonymously (instructors and TAs do not see your name upfront when they mark). What is your preference for anonymous marking?
  - a. I prefer this anonymous marking
  - b. I do not have an opinion on anonymous marking
  - c. I prefer markers clearly know which exams and assignments are mine
4. I felt my overall experience using Gradescope in this course was:
  - a. Very positive
  - b. Somewhat positive
  - c. Neutral
  - d. Somewhat negative
  - e. Very negative
5. What did you like and/or not like about using Gradescope for this course?
6. Is there any other feedback you'd like to provide about Gradescope?

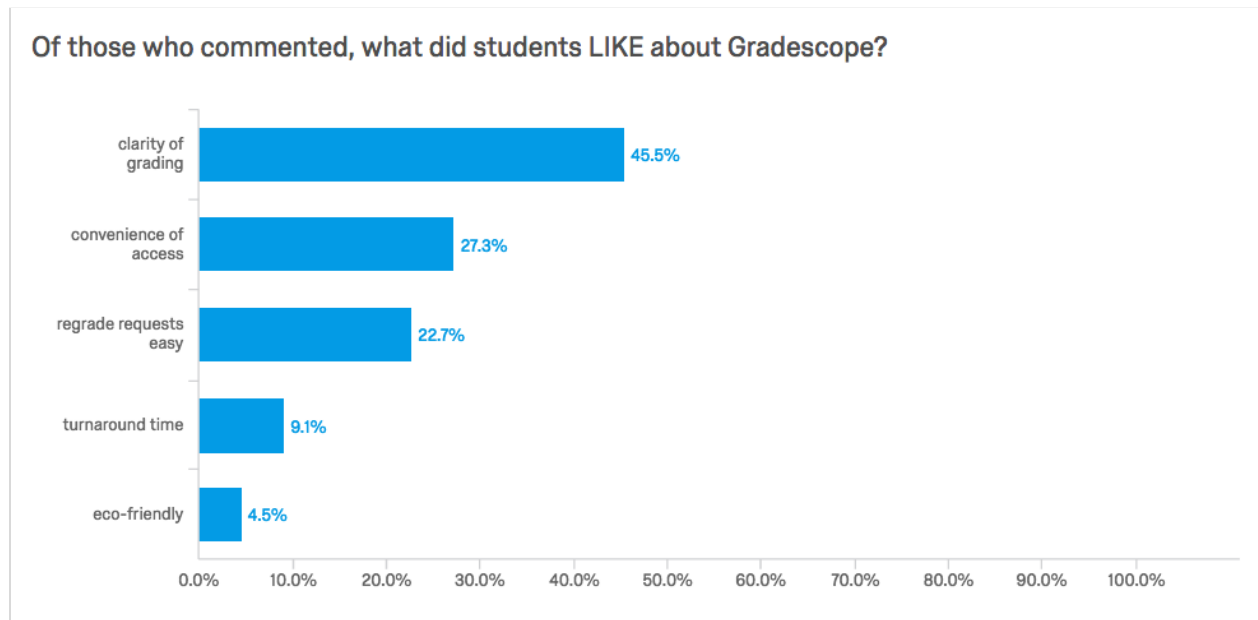


---

## Appendix B - Compiled Student Data

(See: [Qualitative codes](#) for these questions)

### B.1) Themes - Positive Gradescope Student Comments (N = 22)



### B.2) Themes - Negative Gradescope Student Comments (N = 22)

