

OPEN SOURCE WEB-BASED GRADING ASSISTANT PROGRAM: SAHOO EASYGRADE

Debashis SAHOO, PhD
University of California (UC) San Diego.

Megan BARDOLPH, PhD Candidate
University of California (UC) San Diego.

Sheena GHANBARI, EdD
University of California (UC) San Diego.

Jace HARGIS, PhD
NYU Shanghai

ABSTRACT

We examine the application of an online auto-grading system, Sahoo EasyGrade (SEG). SEG is a free, open source grading software that can be used with a simple rubric to grade exams and homework assignments by instructors and teaching assistants. This grading tool was created to facilitate grading exams for a large computer science course where grading individual papers was time-consuming, especially when individual questions required regrades. A survey of the grading team showed that the software made grading more transparent, more fair, and saved time. The survey results were comparable to results of a more extensive survey of a widely used, paid auto-grading system. The present system provided a significant improvement compared to grading by hand, making regrades much more efficient and transparent to the students. This software is available at no cost and is recommended for other courses where grading by hand is time consuming and rapid feedback is desirable for students.

Keywords: Assessment, Measurement, Evaluation, Open Educational Resources (OER), Educational Technology, Auto-grading

INTRODUCTION

The study was conducted at a research intensive university with a culture of collaboration in learning and with an overarching mission to disseminate knowledge for the betterment of society. Accordingly, these initiatives lead to a cultivated, diverse, respectful, open community in which innovation can flourish and align our efforts to be a student-centered, research-focused, and service-oriented university. The mission of our university is to transform by educating, generating and disseminating knowledge and creative works, and engaging in public service. The key strategic university goals are as follows:

- Deliver an educational and overall experience that develops students who are capable of solving problems, leading, and innovating in a diverse and interconnected world;
- Cultivating a diverse and inclusive university community that encourages respectful open dialogue, and challenges itself to take bold actions that will ensure learning is accessible and affordable for all;

- Nurturing and supporting a collaborative and interdisciplinary research culture that advances the frontiers of knowledge, shapes new fields, and disseminates discoveries that transform lives;
- Supporting and promoting just and sustainable forms of economic development, shared prosperity, and social and cultural enrichment regionally and globally
- Creating an agile, sustainable, and supportive infrastructure by ensuring a dedication to service, people, and financial stewardship.

In this environment, this paper examines the application of a grading system called Sahoo EasyGrade (SEG). The author, a professor of Computer Science and Engineering, created this system to fulfill a teaching need. Existing platforms like Gradescope were fee based and there was not an alternative to efficiently managing large scale course grading that took student requests for regrading into consideration. SEG was utilized during the Fall 2017 academic quarter.

The target demographics for this tool are schools which have limited access to particular teaching technologies, such as automated assessments. SEG was used for a Computer Science and Engineering course, but can be generalized to other disciplines. It is suitable for any class with a paper based exam and most effective for larger classes, where it becomes time-consuming to adjust grading items for several exams or to address multiple regrade requests.

In the beginning of the quarter before SEG was created, the instructional team put a lot of effort into grading assignments manually. Students drew attention to a confusing question once grading was completed. The grading team decided that the question was indeed confusing, so they should not penalize students, many of whom lost points on the problem. In order to award credit to any students who missed this problem, the graders had to review every single paper to see whether each lost points, which meant that each paper essentially had to be re-graded. The time cost associated with this activity made it clear that a more efficient system was needed, preferably an online grading system that would allow for a flexible rubric and ease of assignment and exam regrades.

In response to the grading challenges of a large class, the instructor created the online grading system, SEG. This paper examines the benefits and challenges of using this free, instructor-created system for online grading.

LITERATURE REVIEW

Assessment, Measurement, Evaluation

The quality of an assessment tool depends on how it is used (Suskie, 2004). Part of an effective assessment process involves the scoring method. Scoring may be based on a rubric, and every grader should score items according to the rubric, consistently and without errors or bias. It is possible that poor student performance is the result of the assessment itself. In this case, items that are poorly written or inappropriately difficult should be revised.

The American Association for Higher Education's Assessment Forum released a document entitled "Principles of Good Practice for Assessing Student Learning" in 1992 that has been widely used and accepted (Alexander, et al., 2012). One relevant principle states that "Assessment makes a difference when it begins with issues of use and illuminates questions that people really care about." The information that assessment provides is most useful when it is related to relevant questions or problems. It should also be credible and useful to students.

Rubrics

In addition to their usefulness in guiding the creation of assessments and helping students understand learning targets, rubrics help make grading more accurate and more fair (Wolf & Stevens, 2007). When instructors grade individual assignment or exam questions using a rubric,

students can see precisely what mistake(s) they made and how many points were awarded or lost. Instructors are also more likely to grade objectively, assigning the same reward or penalty to every student with minimal bias.

Open Educational Resources (OER)

A U.S. survey of over 2,700 faculty found that only 30% of faculty report awareness of OER. About one half of faculty reported difficulty finding materials and lack of resources for their subject (Seaman & Seaman, 2017). Despite these challenges, awareness of OER and adoption of open-licensed textbooks has been increasing, especially for courses with high enrollment.

Despite advances in OER standards in the U.S. and increased awareness and adoption, many educators have not adopted OER in their classrooms. University professors have reported, among other barriers, difficulty in finding appropriate materials (Berger, 2018). A central location for these resources does not yet exist, and creating such a repository would be a burden for educators.

The OER Research Hub, based in the United Kingdom (UK), surveyed thousands of students and instructors to assess a variety of hypotheses about OER (Weller, Farrow, & McAndrew, 2015). In addition to findings that OER increase student performance and engagement, the researchers found that educators valued the availability of resources, and many reported creating their own materials when suitable resources could not be found. Use of OER also leads to increased reflection for many instructors, leading them to compare their teaching with others', use OER to develop their courses, and to contribute to online discussion of open resources.

Appropriate, Relevant, and Meaningful (ARM) Educational Technology

Although the use of educational technology has seen an increase in higher education classrooms, the search for low threshold, easy to use technology that targets specific needs of faculty and students is still lacking. Through extensive work with faculty development, Hargis (2014; 2010) has identified key attributes for successful technology integration, which includes faculty identifying the appropriate, relevant and meaningful use. Technology can change the nature of a task not just by replacing one method with another (for example, directly substituting online grading for grading by hand), but by allowing new capabilities (Puentedura, 2012). When online grading results are correctly coded and stored, it is possible to use a bank of comments to quickly make notes about common errors, for example. It is also possible to revise grades on individual items, for example, if they are deemed too difficult or poorly worded.

Auto-grading

According to a study of University professors in Australia, instructors value technology that makes assessments easier to perform and easier to grade (Bennett, et al., 2017). Saving time is valuable to the university and to instructors. In an effort to save time, especially in large classes, instructors favored the use of online quizzes and online assignment submission. However, this led to some difficulties such as inefficiency of submitting assignments in paper and online; errors with the design of online discussion; and file submission issues.

Instructors prefer to use modern, innovative technology for assessments but do not always feel that they have the time to focus on technology over teaching (Bennett, et al., 2017). They also find challenges with lack of support for technology-assisted assessment. Cost is also a factor that steers some instructors away from using this technology.

Gradescope is an online system, founded in 2014, that enables semi-automated online grading of handwritten exams and assignments (Singh, et al., 2017). Written answers can be scanned, then graded by a team of instructors and/or teaching assistants and disseminated to students. The creators of this system emphasize its speed, consistency, and flexibility compared to other systems. Rubrics can be created and modified as graders encounter new student answers. These

rubrics allow for greater consistency and fairness in grading. A survey conducted by Singh and colleagues revealed that instructors and teaching assistants believe that Gradescope enhances fairness, saves time, and makes grading more enjoyable.

Methods

In this section we present the participants who integrated the auto-grading system, SEG. This study was conducted at a large research-intensive public university in the southwestern part of the United States. The university supports a Learning Management System (LMS) that is used by many instructors to make course content available to students, track enrollment, and display assignment grades. The LMS can also be used for assignment turn-in, online quizzes, and discussion forums, although some instructors find these features difficult or inconvenient to use. More commonly, instructors use a variety of technologies to supplement the university-supported LMS. The course instructor in the present study desires an LMS that can easily integrate with an online grading system. SEG was envisioned as a first step toward an integrated LMS compatible with large Computer Science and Engineering classes.

Participants

In addition to the course instructor there were graduate and undergraduate students that supported the instructor and also utilized SEG. The instructor had three teaching assistants (TAs) helped to set up questions, lead discussion questions, assist in teaching, lead discussion groups, etc. Additionally, there were 21 paid, trained Tutors, who were peer students, who had previously completed the course. There was no formal training process for TA, the instructor simply provided the tool and actual student papers to grade. TAs graded, asked questions, instructor answered and at times updated the SEG to best suit TA needs.

Through this interface students can take paper exams in real time, leave them in class, the TA would scan and grade using SEG. The process takes about one week and once this is completed a link to the graded test is published on the course LMS. Scores are then returned to students within the same timeframe that the instructor typically operates, but with SEG students are able to receive valuable and detailed feedback within a timely fashion. Students are able to see their paper with marks on it, therefore seeing what they missed and also allowing them to comment on their grading criteria. Furthermore, In SEG students can also request a regrade. This way rather than catering to individual students who request a regrade SEG allows for all students to view the rubric and request a regrade in a manner. In turn, this allows for the TAs and tutors to have process in place to address student concerns while maintaining academic integrity and maintaining a scanned record of examinations.

In addition to creating a streamlined process for grading, it was central to the creator of SEG to make this program accessible and open source. Individuals are able to utilize the system and also implement modifications to improve SEG. A detailed description of how to access SEG is provided in the following section.

Description of technology

Instructors can access SEG using [GitHub](#). Each grader can create a unique login with a password. Exams and assignments can be added to the list of content, which appears on the main page of SEG. Individual exams/assignments can be accessed by selecting from the dropdown menu on the

main page. An example of this page is shown in Figure 1.

Grading Exam Papers

Tools

List of Exams:

1. CSE100 Midterm 1 (+)

- Version A Pages: 2
- Version B Pages: 2
- Version C Pages: 2

CSE100 Midterm 1 ▾

Figure 1: Screen capture of the main page of SEG. This example shows one midterm with three versions. Once logged in, the grader can select content to grade or to view grades.

To grade a specific assignment, the grader can click on the corresponding Select button from the main page. This opens a new page with a drop-down menu to View, Grade, or Match. Each problem number has a specified region on the answer key with a corresponding label. Every assignment displays the specified region in a blue box that moves depending on which problem number is selected in the grading tool. Graders can maintain the selected problem number and view each exam by pressing the Next button at the top of the page, or they can view each

problem for the selected exam using the Prev and Next buttons at the side. Each problem can be graded by selected a button to award full, partial, or no credit. There is a text box to enter comments.

The screenshot shows a scanned exam paper titled "CSE 100 Midterm 1 Answer Sheet Winter 2017 Version A". The paper contains two problems. Problem 1.1 (a) has five sub-questions: (a) True, (b) False, (c) True, (d) False, and (e) False. A blue box highlights the region for question 1.1(b). Problem 1.2 (a) is a diagram with nodes labeled "saoFeng", "worf", and two "mandarin: toki:" boxes. The grading tool overlay on the right includes a name field (Name: Unknown), a score field (0.000/0.000), a dropdown menu (1.1 (b)), and buttons for "Prev" and "Next". Below these are grading options: "1 x 0 Correct: F" and "1 x -1 Incorrect or blank". There is also a "Regrade:" button and a text box for comments. Handwritten red annotations on the paper include a circled "27/29" and "SG".

Figure 2. Screen capture of the grading interface. The right dropdown menu shows Question 1.1(b) selected. The blue box shows the selected Question region. Also shown are the grading options on the right side of the page, Correct (F) and Incorrect or blank; along with a space for optional comments.

Students can request assignment regrades. Graders can view assignments where students have requested regrades by clicking the Regrade button from the main page.

SURVEY

In addition to examining SEG, student feedback was gathered from students who used SEG in CSE100: Advanced Data Structure course in the Fall 2017 quarter. Survey questions were adapted from a published instrument used to assess Gradescope. In the context of assessing Gradescope

derived that speed, consistency, and flexibility were the three greatest strengths of the online grading tool. (Singh et. al, 2017).

The complete list of SEG survey questions are shared below.

Please let us know if you were a TA or Tutor. Rank the following from 1 (strongly disagree) to 4 (strongly agree):

1. Does the system help you grade more fairly?
2. Does the system save you time in grading?
3. Does the system make grading more enjoyable?
4. Does the system simplify regrade requests?
5. Does the displayed rubric help your students learn more from their mistakes?
6. Does the system offer transparency to your students about the grading scheme?

Please read each question carefully and then share a brief response as it relates to Sahoo EasyGrade (SEG).

1. What are the major advantages?
2. What else would you like SEG to be able to do?
3. What are the first three steps in using SEG?
4. What surprised you most about SEG?
5. What attributes of SEG did you find most helpful?
6. What suggestions for change would you have?
7. Please share any other comments.

FINDINGS

Five undergraduate course tutors who used SEG in a Computer Science course completed the online survey. Their responses revealed that saving time and additional transparency of the grading schema were the two compelling benefits of integrating SEG. Additionally, the respondents noted that beyond the speed of grading with SEG it still allows graders to work at their own pace with clear instructions, "it reminds me about the grading scheme for a question at every point," stated a course tutor. Another feature that SEG users reiterated as helpful is the ability to grade and search by question. In general the survey respondents described SEG as "efficient" and "straightforward," but when asked about areas for improvement the findings consistently conveyed a desire for a better user interface.

When asked if SEG helped to grade more fairly all tutors responded that it was an improvement all participants acknowledge that the tool assisted them in grading more fairly, with one hundred percent agreeing (60% agreeing and 40% strongly agreeing) with the statement. Similarly, all of the survey participants found that SEG provided increased transparency, with 40% agreeing and 60% strongly agreeing that the system offered transparency to students about the grading schema. Equally prevalent, all of the participants agreed that SEG saved them time in grading. Sixty percent of participants strongly agreed that the system saved them time while the remaining 40% agreed with this statement.

Additional survey questions were asked to understand the ease of using SEG, benefits of utilizing the displayed rubric. When asked if the system simplified regrade requests 60% of participants agreed, 20% strongly agreed, while one of the participants, 20%, disagreed with this statement. However, all of the participants that that SEG makes grading more enjoyable, with 80% agreeing and 20% strongly agreeing. In regards to rubric usage, all participants note that this feature helps students learn from their mistakes, with 80% agreeing with this statement and 20% strongly agreeing.

Responses to open ended questions also provided additional information about the usability of SEG. In response to the question, "What surprised you most about SEG," 60% of participants stated

that they were surprised that SEG was created from scratch by one person. Additionally, student participants also commented on the ease of grading by question. In response to the question, “What attributes of SEG did you find more helpful,” 80% of participants shared something in respect to the ease of grading by question or navigating between questions.

DISCUSSION AND CONCLUSION

Based on the findings from the survey, speed, transparency, and fairness were the most prevalent improvements to grading. When looking at SEG compared to the assessment of Gradescope, increased speed was the most notable overlap in respective findings. Given the potentially tedious nature of grading in large computer science courses both SEG and Gradescope are tools that save time. As participants stated, SEG is “faster than grading by hand” and “allows me to grade at my own pace and time.”

Although it is difficult to compare data quantitatively to the results of Singh, et al. (2017), the overall trends seem similar. About 60% of respondents strongly agreed that Gradescope saved time in grading, and 26% agreed. There was also a similar response to the question of fairness, with 80% of respondents either agreeing or strongly agreeing that Gradescope helps them grade more fairly.

Results for the question of simplifying regrade requests are also comparable:

- About 74% of respondents agreed or strongly agreed that Gradescope simplified regrade requests, with 15% disagreeing or strongly disagreeing; and
- Twenty percent of SEG survey respondents also disagreed.

Despite both systems’ purported ease and simplification of regrade requests, it seems that not all users have the same experience with processing regrades. This may be an interesting area to explore. Traditionally, regrade requests might be handled individually, with each student who requests a regrade on a single question receiving attention and potential points back. A larger problem arises when many students report that a question was unclear or not fairly graded. In this case, graders may want to adjust all student grades for a single question, and an online system with an existing rubric would be extremely useful in this scenario. Rather than regrading every exam/assignment depending on which students had missed the problem, or uniformly awarding extra points regardless of performance on the question, graders could quickly adjust a rubric, applying that adjustment to every student’s assignment, or they could simply access the single question in SEG and regrade this question without needing to re-examine paper assignments.

We attempted to create a parallel system between the SEG and traditional ways in which students complete their exams and subsequently receive their grades. We found that the instructional team saved time using SEG and overall had a positive experience using this tool. Students were impressed that one instructor was able to create such a functional grading system in so little time. Additional primary benefits of SEG are its accessibility as an open-source software resource and its availability as a free resource.

LIMITATIONS AND FURTHER WORK

Although results from the survey of tutors were relatively consistent, the results represent only a small portion of the grading team and a small number of individuals. There may be information about graders’ experience that are not fully captured.

SEG was only used for one quarter, reducing the generalizability of results, mainly because our university purchased the enterprise version of Gradescope. There is no additional information showing how the system would perform in other classes within the CSE department or in other departments. Instructors who choose to use this tool may discover challenges that were not experienced by the inaugural grading team.

Because our university recently purchased a Gradescope license, which will now perform many of the tasks that the Sahoo EasyGrade (SEG) was able to do and provide a consistent experience for the student, the instructor will use the institutional program. However, the goal of this study is to offer and share the SEG with other schools and universities, who may not have the resources to pay for expensive grading software.

The SEG tool is a web-based application with some python scripts that can access the internal data organization. Therefore, it is extremely versatile and can be integrated other publicly available or privately managed learning management solution (LMS) softwares.

REFERENCES

- Alexander, et al. (2012). Good Practice for Assessment. Am. Association of Higher Ed. Retrieved from www.learningoutcomeassessment.org/PrinciplesofAssessment.html
- Atkinson, R. C., & Shiffrin, R. M. (1971). Human memory: A proposed system and its control processes. *Advances in Research and Theory*, 2, 89-195.
- Bennett, S., Dawson, P., Bearman, M., Molloy, E., & Boud, D. (2017). How technology shapes assessment design: Findings from a study of university teachers. *British Journal of Educational Technology*, 48(2), 672-682.
- Berger, T. (2018) The uncertain future of OER. Retrieved June 4, 2018, from <https://www.edutopia.org/article/uncertain-future-oer>
- Hargis, J. (2014). A ten-year study of faculty classroom observations. *Transformative Dialogues: Teaching and Learning Journal*, 7(2), 1-21.
- Hargis, J. (2010). An online learning environment. *International Journal on New Trends in Education and Their Implications*, 1(3), vi-vii.
- PuenteDura, R.R. (2012). The SAMR model: Background and exemplars. Retrieved from <http://www.hippasus.com/rrpweblog/archives/000073.html>
- Seaman, J., & Seaman, J. (2017). Opening the textbook: Educational resources in US higher education, 2017. Retrieved from <http://www.onlinelearningsurvey.com/reports/openingthetextbook2017.pdf>
- Singh, A., Karayev, S., Gutowski, K., & Abbeel, P. (2017, April). Gradescope: a Fast, Flexible, and Fair System for Scalable Assessment of Handwritten Work. In *Proceedings of the Fourth (2017) ACM Conference on Learning@ Scale* (pp. 81-88). ACM.
- Suskie, L. (2004). *Assessing student learning: A common sense approach*, Boston, MA: Anker.
- Weller, M., De Los Arcos, B., Farrow, R., Pitt, B., & McAndrew, P. (2015). The impact of OER on teaching and learning practice. *Open Praxis*, 7(4), 351-361.
- Wolf, K., & Stevens, E. (2007). The Role of Rubrics in Advancing and Assessing Student Learning. *Journal of Effective Teaching*, 7(1), 3-14.

BIODATA and CONTACT ADDRESSES of the AUTHOR



Debashis SAHOO, PhD

Dr. Sahoo is currently an Assistant Professor in the Department of Pediatrics and Department of Computer Science and Engineering at the University of California (UC) San Diego. He has earned his PhD in electrical engineering from Stanford University.



Megan BARDOLPH, PhD Candidate

Megan Bardolph is a graduate student researcher at the Center for Teaching at the UC San Diego. She is a PhD candidate in the Cognitive Science department at UC San Diego. Her research examines the role of information and affect in attitude change.



Sheena GHANBARI, EdD

Dr. Ghanbari is the Associate Director of the Center for Teaching at the UC San Diego. She completed her Master's in Arts Management from Carnegie Mellon University and her Doctorate in Educational Leadership from UC San Diego and CSUSM. Her research includes creativity and learning and faculty development.



Jace HARGIS, PhD

Dr. Jace Hargis is a Professor and Director of the Center for Teaching & Learning at NYU Shanghai, China. He has authored a textbook, an anthology and published over 140 academic articles. He has earned an MS in Environmental Engineering Sciences and a PhD in Science Education from the University of Florida. His research explores integrating meaningful instructional technologies.