Paper ID #26769

Using Paper-based, Near-immediate Feedback to Support Active Learning in an Introductory Programming Course

Dr. Stewart Thomas, Bucknell University

Stewart Thomas is an Assistant Professor in the Department of Electrical and Computer Engineering at Bucknell University in Lewisburg, Pennsylvania. He received the B.S. and M.Eng. in Electrical Engineering from the University of Louisville in Louisville, KY. and the Ph.D. in Electrical and Computer Engineering from Duke University in Durham, North Carolina. He is a member of ASEE and IEEE.

Using paper-based near-immediate feedback to support active learning in an Introductory Programming course

Abstract

This research paper presents an automated, randomized grading system to support active learning activities and studies its effectiveness in second year (i.e., sophomore-level) engineering programming course. Formative assessment is an essential tool to both the instructor as well as the learner to measure course and concept progression. Further, immediate feedback is known to facilitate student learning, but is not widely adopted. As more classes convert to flipped-style modalities and dedicate significant portions of time to active learning activities, it is important to ensure that students are coming to class prepared. This paper addresses a method to increase opportunities for assessment that includes rapid feedback to the learner.

Online quizzes are one example of a method for providing immediate feedback and gauging student preparedness. However, such a system has its own set of limitations. This paper presents a system to generate randomized assessments for each student from a bank of questions that can be quickly graded and returned to students. The system supports automatic grading of multiple-choice questions and tallying open-response questions marked by the instructor. After scanning in the completed assessments, students are emailed their result with the marked and corrected version as a file attachment.

To evaluate this system, a second year programming course at a small, undergraduate-only midwestern institution was modified to include brief quizzes on daily readings in order to increase time for in-class assignments. The quizzes were distributed during the first few minutes of each class meeting, and students were typically emailed their results within an hour after the completion of the lecture. This system was also used for the two in-class examinations and final examination. Results from the exams indicate that the group of “lost” students constituting the bottom 10% of the class, were supported and brought into the passing grade range. Results from the final exam show a strong improvement of the mean from 77% to 85%. This paper compares exam performance and student-self evaluations between the courses. A discussion of the system and how it facilitates near-immediate feedback and supports flipped classrooms or specifications-based grading schemes concludes the paper.

Background

Creating and using formative assessments effectively in undergraduate courses can be difficult. It is also recognized as essential to the learning process. It is through the process of attempting an activity, receiving feedback, evaluating our attempt, and re-attempting that we learn. As an example, the 2016 video game *The Witness* used rapid feedback to teach players how to solve increasingly complex puzzles using nothing more than a 'pass' or 'fail' response [1], [2]. The game's culmination was a final timed section that involved solving several different archetypes of puzzles the player had previously learned throughout the game. Taking a step back, this begins to sound like the structure of a typical college course — A series of lectures or activities to teach various topics, followed by formative assessment opportunities with feedback culminating in a cumulative, timed final examination.

The challenge to provide students both timely and useful feedback when teaching large-enrollment courses, multi-sectioned courses, or simply multiple courses/preps per semester is increased. Often for the instructor, 'grading' ends up feeling like a balancing act and point of stress, rather than an essential opportunity to provide feedback to the learner. From a time perspective, increasing opportunities for student assessment is not generally seen as a welcome idea even though the research show that the importance of timely feedback [3], [4]. To combat this feeling, and increase overall classroom learning, a variety of technology aids have been explored. Examples include the immediate feedback assessment technique (IFAT) [5], 'clicker' problems [6], or online electronic quizzes through course management software [7].

Though some techniques such as IFAT simplify grading, the grading process is not fully automated. Electronic quizzing offers automatic grading, but students tend to prefer paper-based formats over electronic-based formats — possibly from being more relaxed and less nervous about traditional 'analog' formatting [8]. Even small changes in assessment forms can affect student course evaluation responses. A recent study found use of IFAT forms has a positive impact on course evaluations when compared to traditional Scantron forms [9]. This is especially noteworthy for pre-tenure or other faculty in positions where course evaluations are of high importance. While online quizzing reduces the time required for grading, it may introduce avenues for potential academic dishonesty and has spawned research efforts to investigate methods to minimize 'cheating' in online course offerings [10], [11]. It is also worthwhile to consider the bias that computer requirements in the classroom may bring. Data show that P-12 student performance on computerized testing was positively correlated with socioeconomic status and at-home computer usage [12]. Taken together, what we do in the classroom matters and traditional paper-based assessments remain an important, useful tool for formative assessment. Students value this feedback and are able to gauge their individual learning progression. Instructors are able to gauge the overall progression of the class.

As varying teaching modalities are adopted (e.g., flipped classrooms, contract or specifications grading, etc.), there exists a further need to reduce time spent grading assignments. Even simple assessments such as manually grading brief daily quizzes for 60 students, can take a significant amount of a professor's time. At the same time, implementing short quizzes is an effective approach to enhance active learning [13]. This paper reports on the "automultiplechoice" (AMC) system for automated grading of paper-based exams/coursework. The benefits and tradeoffs of this system are presented and compared to other technologies along with an example of using the AMC system to support an introductory programming course.

Overview of simplified grading technologies

This section provides a brief overview of grading technologies used to simplify or reduce the grading burden of assessments: Scantron forms, IFAT forms, and in-class clickers. Perhaps the simplest grading system is the use of Scantron forms. Scantron systems are used to grade multiple choice questions with students shading in a bubble corresponding to their answer. The student only needs to turn in the 'bubble-sheet' for their work to be graded. This system is limited to grading multiple-choice style questions and the questions provided to students typically only consist of a small number of versions for a single assessment (i.e., many students share the same assignment/form).

The IFAT system is a more recent system with a goal to provide immediate feedback for students. The form consists of a bubble sheet similar in appearance to Scantron forms. Students respond to multiple-choice questions by scratching off a material layer to reveal what is underneath in a manner similar to lottery scratch-off tickets. Students are able to immediately see if their answer is correct (denoted by a star) or incorrect (denoted by a blank). These forms are limited to multiple choice style questions. Typically IFAT forms are graded by awarding partial credit if students scratch off multiple answers. If a revealed response is incorrect, learners can adjust and re-attempt the problem. The IFAT system has been used in many classrooms and been shown to reduce student anxiety in high-pressure testing situations [14].

However, IFAT forms and truly immediate feedback's impact on student learning is called into question based on research in the cognitive sciences. Recent work shows that immediate feedback may not be helpful in long-term retention. In an online quiz, simply altering the delay between between entering a response and viewing the correct answer by four to twelve seconds was found to improve long term retention when compared to immediately viewing the correct answer [15]. It is suspected that this delay activates metacognitive thinking in which learners re-examine their response in anticipation of viewing the answer. Another study found that delaying feedback by a week (as opposed to immediate feedback) also improved long-term retention [15], [16]. Here it was hypothesized that not only the delay, but that student motivation to view their feedback was at play. In either case, immediate feedback such as that available from the paper-based IFAT or from online quizzing, may not be the most effective technique with regard to long term retention.

Finally, using electronic clicker devices is another simple method to assess learning. Once again, this technology is mostly limited to multiple choice style questions. Most clicker systems also require additional technology to be purchased. Of the systems mentioned in this paper, clicker questions are perhaps the most useful for students to see how they are performing in relation to their peers as frequency of collected responses is often revealed to the class.

While there are many options available to aid professors in simplifying grading assessments, most systems are limited to one style of question (multiple choice) and are not in the natural paper-based format of traditional assessments and problem solving exercises. Most attempts to create paper-based automated grading systems work effectively, but tend to be tightly customized to a single professor's preference or for single departmental usage. As an example, a python-based system was developed to support grading paper assignments in a large enrollment course [17]. Data show students perceived the summative and formative assessments using this system to be fair and found the overall experience favorable.

The AMC system is another example a paper-based, automated grading system. It was selected for its general usage capabilities and ability to provide timely feedback to the learning in an electronic format. This system is further described in the next section.

Paper-based automatic grading using AMC

The AMC system is an open-source software package available for download from its main website [18]. It was originally created in 2008 and remains under active development. The system combines several open-source technologies. A significant advantage of assignments (or exams) created in AMC is that questions are not solely limited to multiple-choice style questions. The system can also create and tally open-response questions in which students could be asked to to draw a diagram, correct a given image, write a prose response, fill in the blank, etc. The major features that AMC offers are:

- Each individual student's form/exam is unique — *every student has a different copy!*
- Quickly provides formative assessment of course progression — Students see their progression and professors see the progression of the learners as a whole
- Allows for hand-written feedback to the learner
- Randomized ordering of questions (per assessment)
- Randomized ordering of responses to multiple choice questions (per assessment)
- Randomized selection of questions from a larger question bank
- Advanced grading schemes (such as allowing students to skip a question with no penalty)
- Ability to grade/tally open-response questions (requires professor input)
- Use random numbers to generate different values/answers for every instance of a problem
- Analog, paper-based exams
- Send students graded and corrected version of *their* exam via email
- Retain graded assessments on file, electronically
- Electronically tally grades into a spreadsheet

The process for creating an assessment using AMC is shown in Fig. 1. First, create a bank of questions. These questions can either be created in a plain text format using AMC's custom markup formatting, or using the LaTeX language. A 'build' is then created which lists/selects the questions to use or describes how to randomly draw the questions. The AMC system then creates a unique form for each student. These are printed and handed out. Students bubble in answers to

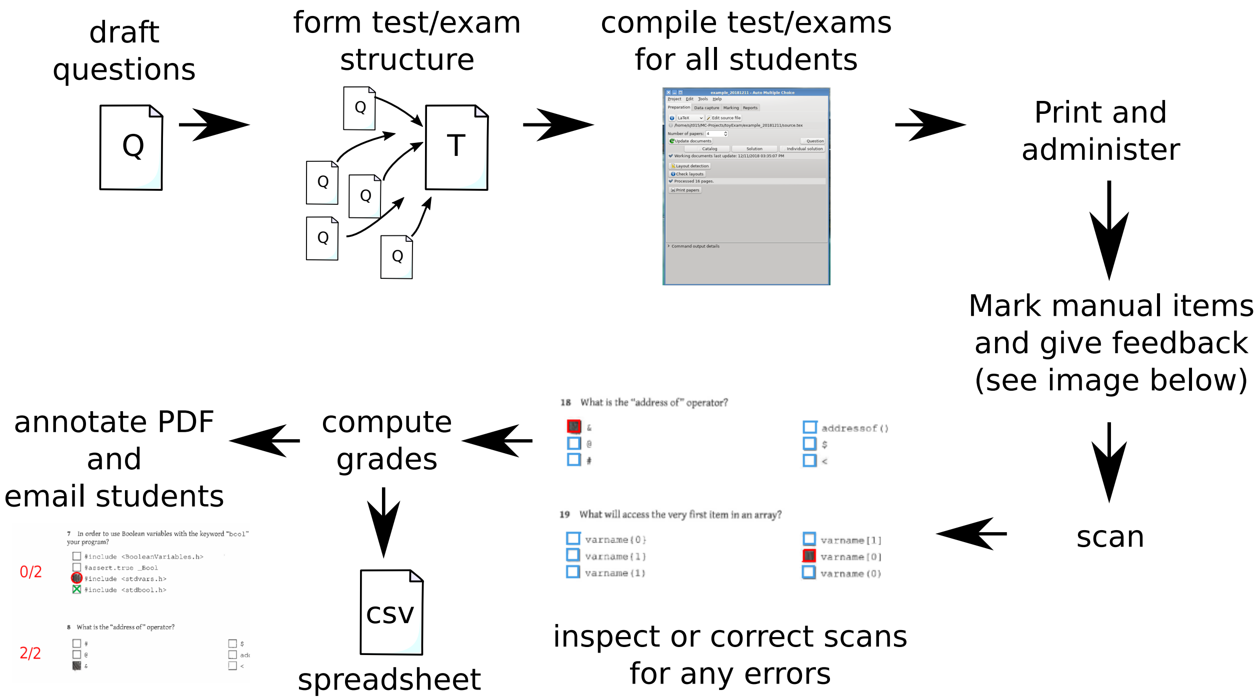

Figure 1: Grading process in AMC from creating questions to returning grades.

multiple choice questions, and answer open-response questions as they normally would on a traditional assessment. The professor collects and grades the open problems by marking one of the appropriate boxes (e.g., full-credit box, partial-credit box(es), no-credit box). This step is performed on paper and allows for the professor to provide handwritten feedback to the learner in the context of their response. Once complete, the paper assessments are then scanned (using an office copier) and loaded into the AMC program. The AMC program will automatically tally the individual marks and note correct/incorrect responses. Usually at this step, the professor reviews the automated marking to check for errors (i.e., students checked multiple boxes, changed answer, etc.) and make adjustments as needed. After reviewing the scans for mistakes, the grades are electronically tallied and associated with a student name. This name-to-form association can be done by manually viewing a scanned image of the students' handwritten name, or automatically by using student ID numbers. Each scanned, tallied form is then converted to a portable document format (PDF) and annotated to include the student's overall score and scores for each individual question. For questions answered incorrectly, the correct answer is clearly marked alongside the student's (incorrect) answer. These annotated files are then emailed to the student's email address. Samples of graded exams that the students would receive in email are shown in Fig. 2. In this figure (a) shows the overall grade, (b) show an example of a correct response, (c) shows a graded open-response question with included instructor feedback, and (d) shows an incorrect response that includes the correst response marked. This feature is especially beneficial for providing students formative assessment information.

A key benefit of this system is that a professor can administer assessments unique to *each student in a course* and quickly return the graded, corrected forms as an email attachment. Providing students individually unique assessments is a simple method to help minimize the opportunities

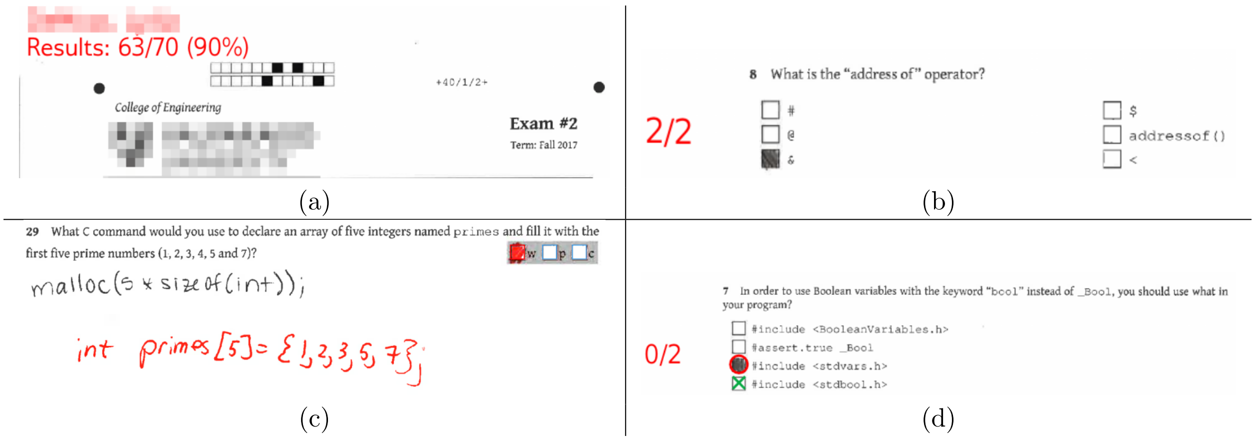

Figure 2: Sample of a corrected exam sent to students as an email attachment showing (a) first page with total score (anonymized), (b) sample of a correct student response, (c) an open-response question with handwritten feedback, and d) an incorrect student response showing.

for academic dishonesty in a classroom, though this typically requires a great deal of effort. With AMC, such uniqueness is automatic. Because each exam is filed electronically, it is also a simple matter to address student requests for re-grading — even with the many versions of assessments.

The AMC system was recently implemented in an introductory programming class. In this class, quizzes on daily readings were introduced in order to allow more time for students to work on in-class exercises. To reduce the grading burden, AMC was used to create, grade and return the daily quizzes. As a user of the system, there were several secondary features that were found to be very helpful:

- Easy grading of 45 daily quizzes throughout the semester and always returned to students the same day — *45 assessment opportunities were added to this course*
- Building up a question bank from daily quizzes *significantly* reduced the time needed to prepare exams
- Retention of scanned copies from *all* reading quizzes and course exams (useful for accreditation record keeping)
- Because exam results are searchable in an email archive, a phone (or other computer) can be used to respond to students' grading questions immediately (i.e., in-class) and see the student's work.

Experimental Overview

ECE 251 is an introductory programming class taught to all second-year electrical and computer engineering students at a small, undergraduate-only midwestern institution. This course teaches the C programming language and is most students' first course in programming. Course topics include: compilation, variable types, functions, pointers, standard library usage, structures, memory access, file input-output, and brief primer on OpenGL 3D graphics.

In Fall 2016, the AMC system was not used, and the reading quizzes were not implemented. This course met for three 50-minute meetings a week. In a typical lecture, the instructor distributed paper copies of the notes for the day, lectured on the material for about half of the class meeting and then turned over the balance of time to the students to work in-class on the day's assignment in pairs as the professor answered individual questions as they arose. These assignments were due at the start of the next meeting. Enrollment was 60 students across two sections.

To support more time for students to work on course assignments in-class, the Fall 2017 offering used the AMC system to include more assessment opportunities in the form of daily reading quizzes. The Fall 2017 offering met for two 75-minute meetings a week. In a typical lecture, the instructor distributed paper copies of notes for the *next* meeting. Students were expected to read these notes before coming to class. In the first five-minutes of every class meeting, students completed a brief quiz typically consisting of two or three questions. While students' individual quizzes consisted of 2 – 3 questions, a bank of approximately 6 – 10 questions for each lesson was created. After the quiz, the professor briefly discussed the material (typically 15 – 20 minutes, although this varied by lecture) and highlighted or provided additional details to the reading. Students spent the remainder of time working on the day's assignment in pairs while the professor provided feedback and answered questions to individual groups. Assignments were due at the start of the next meeting. Enrollment was 45 students across two sections.

To gauge the effectiveness of the paper-based randomized quizzing, the results from two mid-semester exams are compared, along with final exam scores between semesters. Data from student end of semester course evaluations showing student self-assessment of the course topics are also compared. Finally, a summary of the professor's experience is shared.

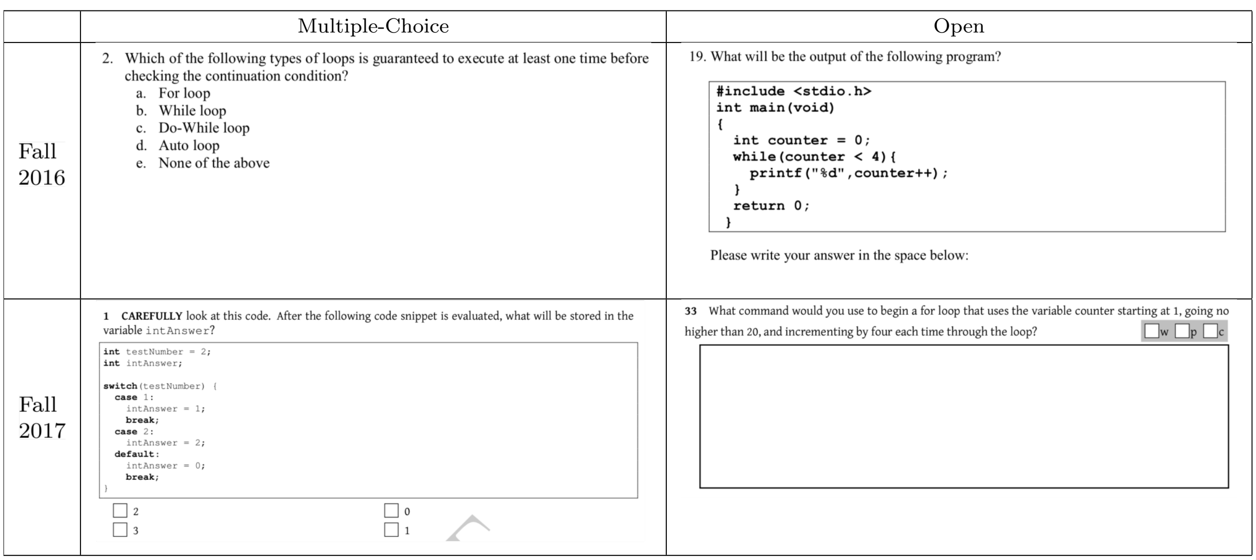

Figure 3: Sample questions from Exam 1 showing similar appearance between multiple-choice style and open style questions between Fall 2016 and Fall 2017.

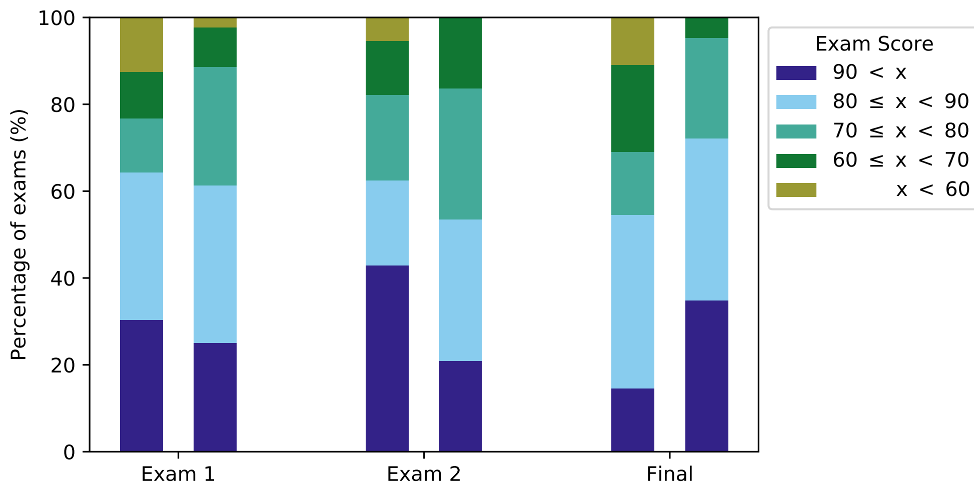

Figure 4: Stacked bar graph of exam results by score. Scores from 2016 are the left bars, scores from 2017 are the right bars.

In ECE 251, three exams (two mid-semester exams, and one cumulative final) were administered in both the 2016 and 2017 offerings. The first exams consisted of an in-class portion and a take-home programming assignment. The final exam was cumulative and had no take-home portion, although students had previously completed a large, multi-week final project. Though difficult to directly compare, in both offerings performance on the final projects were comparable with most students meeting or exceeding expectations.

Results

As Fig. 3 show, formatting of exams across the Fall 2016 offering without daily quizzes and the Fall 2017 offering that included daily quizzes was similar. In 2016, the in-class portion of the first exam had 35 questions that varied between multiple-choice style questions and open-responses questions. Similarly, the Fall 2017 offering's first exam had 47 questions that varied between multiple-choice style questions and open-response questions. These questions were a mix of questions from the bank of daily quiz questions and questions developed specifically for the exam.

A breakdown of results for all exams is shown in Fig. 4. Score distributions for each exam is shown as a stacked bar plot. Exam scores above 90% are shaded dark blue and represented as the bottom-most segment of the bar graphs. Scores between 80% and 90% are shaded light blue and stacked on top of the dark blue block, etc. Each exam is represented by a pair of bar plots with scores from 2016 offering that did not use daily quizzes on the left, and the scores from the 2017 offering that implemented daily quizzes on the right. Only exam scores from the in-class portion are shown.

Though it is difficult to directly compare exam scores from different offerings, there are two trends to note. First, in comparing the final exams between offerings, the mean score in 2016 was

Table 1: Course Evaluation Questions

#	Question ("I can . . . ")	$\sum r \geq 3$ (%) 2016	2017
1	Write a computer program to perform scientific calculations.	100.00	100.00
2	Use formatted input and output to interact with the user.	100.00	100.00
3	Use appropriate data types to store different types of variables.	100.00	100.00
4	Use conditional branching statements to implement decision making and program flow control.	100.00	100.00
5	Use definite (for-next) and indefinite (do-while) loops to perform repeated actions.	93.55	100.00
6	Use arrays and strings effectively to solve programming problems.	96.77	100.00
7	Use functions and subroutines to create structured programs	87.10	100.00
8	Use pointers and dynamic memory allocation to solve programming problems	83.87	75.00
9	Read and write data to and from a file.	80.65	87.50

Students responded on a 5-point weighted Likert scale with:
'5' representing 'Yes, definitely' and '1' representing 'No, not at all.'

77.23% compared to 85.63% in 2017. The offering that included daily quizzes showed a higher mean final exam score. Second: in the 2016, 23% of students scored below 70% on the first exam. At the final exam, 30% of the students scored below 70%. Though this group appeared to show improvement between the first and second exam, by the end, their final exam score did not reflect mastery of the course concepts.

However, the trends of the 2017 offering that included daily quizzes indicate a slightly different story. In this case, there were still students that performed poorly on the first exam — 18% of students scored below 70%. By the end of the course, only 5% scored below 70% on the final exam. Most of the students who were struggling the most with the course (the lowest 10%) were able to be pulled up into the passing grade ranges and not 'left behind' as the course progressed.

While exam scores are only one aspect of the course, it is should be noted that there were no significant shifts in performance of final projects or take-home portions of the examinations. Available exam data suggests that the increased opportunities to receive feedback provided by the daily quizzes and the increased time for in-class problem solving were useful for students' progression in the course while not presenting a significant burden for the professor.

Student responses on the end of course evaluations, also support that the daily quizzes contributed to their overall learning. In response to the questions shown in Table 1, students self-reported their ability using a 5-point weighed Likert scale. In this scale, a 5 represented "Yes, Definitely" and a 1 represented "No, not at all." Tallying responses of 3, 4 or 5 to the questions (indicating a neutral or positive response to the question), show that students in the offering with daily quizzes felt more confident in their abilities for all but one item. This is summarized in the rightmost column of the table which displays the sum of responses for 3, 4 and 5 as a percentage of total responses. This data supports similar findings from the course examinations that the daily quizzes and associated feedback were helpful to student learning.

Students in the 2017 offering with reading quizzes were also asked the question "Did the daily quizzes help you learn the material?" Responses were tallied on the same 5-point weighted Likert scale (5=Yes, Definitely, 1=No, Not at all). No students responded to this question with a 1 or a 2. 68% of students responded with a 4 or 5 and the remaining 32% responded with a 3. In the 2017 evaluations, one student noted: "Good class, homework/quizzes/tests were fair."

Discussion

In looking at the results, data suggests that implementing daily quizzes were beneficial to student learning. The number of opportunities for assessment and feedback to the learner was greatly increased while also increasing the amount of time for in-class problem solving. However in this study it is difficult to isolate the effect of the daily quizzes from the change in number of weekly meetings or increase of time for in-class problem solving. Further the course evaluation data remains difficult to interpret. In the 2016, only 51% of students enrolled in the course participated in the final course evaluations and in 2017, only 35% of students completed the evaluations. Because of these low response rates, it is possible that the reported data in this study does not accurately reflect general student perceptions of the course. More work is need to isolate the effects of the daily quiz and to study the effect of the various feedback methods.

From the professor's perspective, the AMC tool is a great help. Even though quiz and question development required a non-trivial amount of time, generating a prepared bank of questions decreased the time taken to prepare examinations. Using the AMC system, grading and returning quizzes required very little effort. This allowed for time to begin preparing the next section/lecture immediately after an exam and not be burdened with grading and preparing a lectures simultaneously.

As a formative assessment tool, AMC provided useful insight into individual student and overall progress throughout the course. For every quiz, grades were automatically compiled into a spreadsheet displaying basic statistics of how questions were answered. Throughout the course, it was a simple matter to check for course concepts students were having trouble understanding. These misunderstandings could be caught early (i.e., before a major exam) and lectures were able to be adjusted to address these issues. This was helpful in addressing student understanding of pointers in C and variable types early in the course. Lectures were altered to further review the concepts, and related questions were again included in the daily reading quizzes. Having a record of historical data and overall progression saved across different course offerings creates an interesting system for professors looking to gauge continual improvement in their courses. Having this baseline will allow for studies such as this one to take place — evaluating the effect of course changes compared to historical trends.

In this study, the AMC system was used to support an increase in time for in-class problem solving. This study shows that the AMC tool is convenient for those looking to support a flipped classroom environment. In terms of contract or specifications grading [19], a major challenge is the creation of a variety of assignments for students. In this learning modality, assignments are graded pass/fail and students are allowed to re-attempt problems until they pass. Using AMC to automatically generate problem sets helps support the conversion to contract grading. Students receive assignments with randomly selected problems that can be quickly graded and returned with feedback.

As mentioned previously, immediate access to a student quiz or exam via a phone (or computer with email access) proved surprisingly useful. Frequently, students would have a minor question about a problem or ask about an question in or between classes. Immediately, I was able to view the graded exercise and address the problem in the context the student's solution. Because of the favorable experience of using AMC system, I have implemented it in all my classes. In a recent conversation, a past student informed me they found having copies of graded assessments available in email to be a great help during the course.

However, there are a few disadvantages of the system that should be noted. In its current state, AMC is not a web-based application, although such a conversion is being investigated [20]. The tool is intended to be run on a Linux computer and to take full advantage of all the features, requires users to have a working knowledge of the LaTeX typesetting system. These are not strict requirements as a Mac (Apple) version is available as well as a non-LaTeX text format for writing questions. Nonetheless, these requirements can be difficult to overcome as it is not (yet) a simple plug-and-play solution. Further, AMC does not directly integrate with learning management software (LMS). For many users, this is actually advantage as the AMC system is completely standalone and not tied to a particular LMS version or vendor. For users of LMS, the grade results and spreadsheet format AMC generates can easily be imported into most learning management software.

Conclusion

The AMC system is a useful tool that allows for automated grading of randomized assessments. Specifically, grading is automated and corrected forms are emailed to students as PDF attachments. The system is extremely flexible and allows for a large variety of assessment formats. While open-response questions are not automatically graded, student papers can be quickly marked by a professor and included in the automatic calculation of grades. Students still receive handwritten, individual feedback from professors, readily available via email.

As a case study, an introductory programming class implemented daily quizzes over assigned readings that were created and graded using AMC. With minimal burden on the professor, the offering that included daily quizzes showed higher exam scores with fewer students self-reporting a low confidence in their abilities. Overall, students were favorable to the automated grading system, perceiving it as a useful, fair system.

In looking forward, the AMC system allows a course to increase opportunities to assess student understanding of concepts and progression in the course while minimizing the grading burden. The created assessments are unique to each individual student, and are stored electronically. In courses implementing specifications-based (sometimes called 'contract') grading, often assignments may be attempted multiple times. This can be difficult to administer when many students are attempting different assignments simultaneously. However, systems such as AMC can simplify such a transition as many unique problems can be generated (when using random numbers) and automatically graded.

References

[1] One More Continue. (May 2017). The Methodology Behind The Witness, [Online]. Available: https://onemorecontinue.com/methodology-behind-the-witness/.

[2] M. Haydon, “Gaming the Education System: Interactive Game-Based Learning”, *All Regis University Theses*, Jan. 2017.

[3] J. Chen, J. Kadlowec, and D. Whittinghill, “Using rapid feedback to enhance student learning”, in *2006 ASEE Annual Conference & Exposition*, Chicago, IL: ASEE, 2006.

[4] R. Higgins, P. Hartley, and A. Skelton, “The conscientious consumer: Reconsidering the role of assessment feedback in student learning”, *J. Studies in Higher Education*, vol. 27, no. 1, pp. 53–64, 2002.

[5] Epstein Educational Enterprises. (2019). “IF-AT by Epstein Educational Enterprises”, [Online]. Available: http://www.epsteineducation.com/. [Accessed Jan. 7, 2019].

[6] S. H. Cotner, B. A. Fall, S. M. Wick, J. D. Walker, and P. M. Baepler, “Rapid Feedback Assessment Methods: Can We Improve Engagement and Preparation for Exams in Large-enrollment Courses?”, *J.Science Education and Technology*, vol. 17, no. 5, pp. 437–443, Oct. 2008.

[7] B. R. Cook and A. Babon, “Active learning through online quizzes: better learning and less (busy) work”, *J.Geography in Higher Education*, vol. 41, no. 1, pp. 24–38, Jan. 2017.

[8] S. Washburn, J. Herman, and R. Stewart, “Evaluation of performance and perceptions of electronic vs paper multiple-choice exams”, *Adv Physiol Educ*, vol. 41, 2017.

[9] T. W. Maurer and J. J. Kropp, “The Impact of the Immediate Feedback Assessment Technique on Course Evaluations”, *Teaching & Learning Inquiry: The ISSOTL Journal*, vol. 3, no. 1, p. 31, 2015.

[10] D. N. Buechler, “Online Quizzing and Incremental Feedback for Distance and Local Students”, in *2017 ASEE Annual Conference & Exposition*, ASEE, Jun. 2017.

[11] M. Farrell. (Mar. 2015). Want to stop cheating on online quizzes? . . . . Let them cheat!, [Online]. Available: https://onlinelearningconsortium.org/want-stop-cheating-online-quizzes-let-cheat/. [Accessed Jan. 2019].

[12] J. E. Kilgore Jr., “Exploring the factors that influence attitudes and achievement when students take computerized tests”, PhD thesis, Walden University, 2008.

[13] M. L. Epstein and G. M. Brosvic, “Students Prefer the Immediate Feedback Assessment Technique”, *Psychological Reports*, vol. 90, no. 3, pp. 1136–1138, Jun. 2002.

[14] D. Dibattista and L. Gosse, “Test Anxiety and the Immediate Feedback Assessment Technique”, *J. Experimental Education*, vol. 74, no. 4, pp. 311–328, Jul. 2006.

[15] H. G. Mullet, A. C. Butler, B. Verdin, R. von Borries, and E. J. Marsh, “Delaying feedback promotes transfer of knowledge despite student preferences to receive feedback immediately”, *J.Applied Research in Memory and Cognition*, vol. 3, no. 3, pp. 222–229, Sep. 2014.

[16] K. M. Mullaney, S. K. Carpenter, C. Grotenhuis, and S. Burianek, “Waiting for feedback helps if you want to know the answer: the role of curiosity in the delay-of-feedback benefit”, *J. Memory & Cognition*, vol. 42, no. 8, pp. 1273–1284, Nov. 2014.

[17] G. L. H. Matthew West Mariana Silva, "Randomized Exams for Large STEM Courses Spread via Communities of Practice", in *2015 ASEE Annual Conference and Exposition*, Jun. 2015, pp. 1–26.
[18] A. Bienvenüe. (2019). "AMC – Multiple Choice Questionnaires management with automated marking", [Online]. Available: http://www.auto-multiple-choice.net/. [Accessed Jan. 5, 2019].
[19] J. Mendez, "Standards-Based Specifications Grading in a Hybrid Course", in *ASEE IL-IN Section Conference*, Jun. 2018.
[20] C. Primozic, J. Roush, R. Schenck, B. Levandowski, B. Doriott, and W. Sullivan, "Interface Extensions to the AMC Evaluation Platform", *Symposium on Undergraduate Research and Creative Expression (SOURCE)*, May 2018, (Valparaiso, IN).