Assessment in Online Learning: It’s a Matter of Time

Charles Kerton  
Iowa State University, kerton@iastate.edu

Cinzia Cervato  
Iowa State University, cinzia@iastate.edu

Follow this and additional works at: http://lib.dr.iastate.edu/ge_at_pubs

Part of the Astrophysics and Astronomy Commons, Earth Sciences Commons, Educational Assessment, Evaluation, and Research Commons, Educational Methods Commons, Higher Education Commons, Instructional Media Design Commons, and the Online and Distance Education Commons

The complete bibliographic information for this item can be found at http://lib.dr.iastate.edu/ge_at_pubs/15. For information on how to cite this item, please visit http://lib.dr.iastate.edu/howtocite.html.
Assessment in Online Learning: It’s a Matter of Time

Abstract
Taking online courses is becoming a more common part of the college experience, but very little is known about student behaviors and strategies related to online assessment. This article reviews how students in an online Earth and Space Science course interact with various online assessments. Our two main findings are that our students do not use self-assessment tools effectively, and time spent on online exams is surprisingly short. We discuss how the use of self-assessment tools can probably be improved through careful online course design, but the short time spent on online exams is partially due to the nature of the online environment itself. We make a number of design suggestions that can encourage good test-taking strategies in the online environment.

Keywords
digital natives, online assessment, self-assessment, test-taking strategies

Disciplines
Astrophysics and Astronomy | Earth Sciences | Educational Assessment, Evaluation, and Research | Educational Methods | Higher Education | Instructional Media Design | Online and Distance Education

Comments
This article is from Journal of College Science Teaching 43 (2014): 20. Posted with permission.
Assessment in Online Learning—It’s a Matter of Time

By Charles Kerton and Cinzia Cervato

Today, “digital natives” (Prensky, 2001) make up the vast majority of college students, and many of them are increasingly testing their digital abilities in distance education courses. The debate on how proficient they are in the use of e-learning technology is still ongoing (e.g., Margarayan, Littlejohn, & Voigt, 2011); nevertheless, little is known about their strategies for online assessment.

We recently designed a two-credit online course in Earth and Space Science for elementary education majors (Cervato, Kerton, Peer, Hassall, & Schmidt, 2013). The course, taken by between 35 and 48 students per semester, with between 64% and 94% of those being first- and second-year students, is divided into 15 weekly content modules consisting of online multimedia content and offline readings, projects, and experiments. The course was delivered using the Blackboard Learn 9 Learning Management System (LMS). Each module was designed to include a suite of online formative and summative assessments (worth 60% of the final grade): multiple-choice self-assessment (SA) questions to help students navigate the associated online and text readings; a nonproctored, open-book, multiple-choice quiz on the module content; and a proctored exam every three or four modules. The remaining 40% of the final grade was based on various semester-long projects.

Over three semesters we experimented with different ways of using the various assessment tools and explored how our students interacted with them. We have found that (a) our students do not use SA tools effectively, and (b) time spent on online exams is surprisingly short.

Use of self-assessments

We implemented a simple style of self-assessment—consisting of 10 questions randomly drawn from a large pool—that students were to take after completing the online module. Students were given unlimited time to complete the questions and an unlimited number of attempts. Completion of each SA quiz with a score of 8 or better was worth 0.3% of their final grade. The goal of the SA was to allow the student to test his or her understanding of the content material without the fear of negative consequences for getting answers incorrect. Our intention was that the students would use the feedback from the SA to identify areas in which they are weak; this would lead to a more efficient use of their study time and ultimately better performance in the course as a whole.

We found that in Semester 1, when use of the SA was required, the students accessed the SA frequently, but there was no correlation between SA usage and a student’s final grade. In Semester 2, there was no credit associated with the activity, and almost none of the students, including the better ones, used the SA quizzes to assist them with the course (see Figures 1 and 2); in Semester 3, we did not use the SA quizzes at all. In Semesters 2 and 3, overall student performance in quizzes and exams was better than in the first semester. This was not expected, as instances where SA usage does have a positive...
Assessment in Online Learning—It’s a Matter of Time

correlation with a student’s grade have been reported in the literature at least for face-to-face courses (Kibble et al., 2011; Smith, 2007).

Why was SA so ineffective in our course, to the point that removing it improved student performance? We believe that the foremost reason is that the ability to use SA effectively requires the ability to reflect on one’s progress in learning, a high-level metacognitive and knowledge skill, using the language of Bloom’s Taxonomy (e.g., Heer, 2009), that many of our students have not yet developed. One way to describe students’ learning progress is as a journey from novice to expert (Heck & Wild, 2011). Expert learners have developed metacognitive skills that allow them to monitor their learning, change strategies, and assess their progress. Novice learners have not—they view metacognitive opportunities such as formative assessment as duties or requirements. As such, our students are novice learners in science and in the online environment.

Another possibility is that the way SA was included in the course, as a separate task following reading, is not an effective way to engage and scaffold learning in this group of students. We will return to this idea after we analyze a second aspect of online assessment: how much time students invest in online testing.

Duration/timing of online assessment

On the basis of our experience in face-to-face courses and best practices for design and administering of multiple-choice exams (e.g., Brothen & Wambach, 2004; Chronicle of Higher Education, 2010), we decided that 1–2 minutes per question would allow sufficient time for all students to complete the exams. Each quiz randomly selects questions from a test bank pool provided by the textbook publisher, and each exam pulls questions from the combined pools of the quizzes that covered the same material. Questions are administered one at a time, backtracking is permitted, and the double-check the answer if needed and at 1 hour for 30-question exams. Each quiz randomly selects questions from a test bank pool provided by the textbook publisher, and each exam pulls questions from the combined pools of the quizzes that covered the same material. Questions are administered one at a time, backtracking is permitted, and the

FIGURE 1
Semester I self-assessment (SA) usage. Each circle shows the number of SA activities completed by students and the average time they spent on each SA. The diameter of the circle is proportional to their final letter grade in the course as shown in the lower right. Note that the majority of students completed all of the 14 required SA activities and that there is no strong correlation between SA usage and final course grade.

![FIGURE 1](image1.png)

FIGURE 2
Semester 2 self-assessment (SA) usage. When not graded, SA usage dropped dramatically. Note that only two students completed all 14 available SA activities and that there is no correlation between SA usage and final course grade.

![FIGURE 2](image2.png)
LMS warns the student before submission if an answer was left blank. After submission, students can see the questions, their answers, and the points they earned for each question. They do not see the correct answer.

To evaluate our approach, each semester we analyzed how much time students spent in each assessment (see Tables 1 and 2). On average, students spent between 10 and 13 minutes on each exam (worth up to 10% of the course grade) and between 13 and 16 minutes on each quiz (worth up to 1% of the course grade). In terms of average time spent per question, this translates to between half a minute per question on exams to one and a half minutes per question on quizzes. In both cases this includes the time the student would spend reading a question, reading up to four multiple-choice answers, evaluating the options, making their selection, and submitting it. For the quizzes this would also include any time the students spent consulting their notes, textbook, or other available resources, likely accounting for the time difference seen between quizzes and exams. We also found no significant correlation between exam or quiz score and time spent on it, and we found no correlation between time spent on quizzes or exams and final course grade.

We were not surprised by the lack of correlation between online evaluation time and final grade; from personal experience with our students, we know that a long exam duration...

---

**TABLE 1**

Quiz summary.

<table>
<thead>
<tr>
<th>Evaluation</th>
<th>Number of questions</th>
<th>Average time* (minutes)</th>
<th>Time per question (seconds)</th>
<th>Average score*</th>
<th>r (time vs. score)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fall 11 quiz</td>
<td>10</td>
<td>15.85 (5.0)</td>
<td>95</td>
<td>7.82 (0.8)</td>
<td>0.46</td>
</tr>
<tr>
<td>Spring 12 quiz</td>
<td>10</td>
<td>13.16 (3.6)</td>
<td>79</td>
<td>7.64 (0.9)</td>
<td>0.42</td>
</tr>
<tr>
<td>Fall 12 quiz</td>
<td>10</td>
<td>15.84 (3.8)</td>
<td>95</td>
<td>8.48 (0.6)</td>
<td>–0.07</td>
</tr>
</tbody>
</table>

*a standard deviation is given in parentheses.

**TABLE 2**

Exam summary.

<table>
<thead>
<tr>
<th>Evaluation</th>
<th>Average total time (minutes)*</th>
<th>Time per question (seconds)</th>
<th>Average score*</th>
<th>r (time vs. score)</th>
<th>Average # of new questions</th>
<th>Adjusted time per question** (seconds)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fall 11 – 1</td>
<td>14.26 (5.9)</td>
<td>28.5</td>
<td>17.35 (3.3)</td>
<td>0.33</td>
<td>23</td>
<td>35.7 (34.2)</td>
</tr>
<tr>
<td>Fall 11 – 2</td>
<td>10.71 (5.3)</td>
<td>21.4</td>
<td>17.55 (4.2)</td>
<td>0.01</td>
<td>18</td>
<td>32.4 (29.0)</td>
</tr>
<tr>
<td>Fall 11 – 3</td>
<td>8.79 (2.7)</td>
<td>17.6</td>
<td>16.26 (3.4)</td>
<td>0.16</td>
<td>15</td>
<td>30.2 (25.2)</td>
</tr>
<tr>
<td>Fall 11 – 4</td>
<td>10.31 (4.3)</td>
<td>20.6</td>
<td>18.77 (3.6)</td>
<td>–0.16</td>
<td>18</td>
<td>31.0 (27.7)</td>
</tr>
<tr>
<td>Spring 12 – 1</td>
<td>13.46 (5.2)</td>
<td>26.9</td>
<td>18.30 (3.8)</td>
<td>–0.10</td>
<td>23</td>
<td>33.6 (32.1)</td>
</tr>
<tr>
<td>Spring 12 – 2</td>
<td>9.70 (3.2)</td>
<td>19.4</td>
<td>20.11 (4.7)</td>
<td>0.03</td>
<td>18</td>
<td>29.0 (25.7)</td>
</tr>
<tr>
<td>Spring 12 – 3</td>
<td>8.67 (3.0)</td>
<td>17.3</td>
<td>20.96 (4.7)</td>
<td>0.08</td>
<td>15</td>
<td>29.7 (24.7)</td>
</tr>
<tr>
<td>Spring 12 – 4</td>
<td>8.33 (2.2)</td>
<td>16.7</td>
<td>22.11 (4.8)</td>
<td>–0.11</td>
<td>18</td>
<td>24.4 (21.1)</td>
</tr>
<tr>
<td>Fall 12 – 1</td>
<td>17.57 (5.8)</td>
<td>35.1</td>
<td>19.60 (2.9)</td>
<td>0.12</td>
<td>23</td>
<td>44.3 (42.8)</td>
</tr>
<tr>
<td>Fall 12 – 2</td>
<td>11.11 (3.6)</td>
<td>22.2</td>
<td>21.88 (4.2)</td>
<td>–0.31</td>
<td>18</td>
<td>33.7 (30.4)</td>
</tr>
<tr>
<td>Fall 12 – 3</td>
<td>11.34 (4.7)</td>
<td>22.7</td>
<td>22.07 (5.0)</td>
<td>–0.41</td>
<td>15</td>
<td>40.4 (35.4)</td>
</tr>
<tr>
<td>Fall 12 – 4</td>
<td>12.45 (4.7)</td>
<td>24.9</td>
<td>22.07 (4.6)</td>
<td>–0.26</td>
<td>18</td>
<td>38.2 (34.8)</td>
</tr>
</tbody>
</table>

*a standard deviation is given in parentheses.

**assumes the student spends 5 (or 10) seconds on each previously viewed question.
Assessment in Online Learning—It’s a Matter of Time

could equally be due to a poor student struggling or a good student being meticulous and rechecking answers, and this intuition has been confirmed by many educational and psychological studies as reported by Schnipke and Scrams (2002). However, we were puzzled by the very short time students spent doing the online assessments and, in particular, how the amount of time spent was not influenced by the amount of course grade linked to each assessment type. We could not find studies that showed how much time students typically spend on multiple-choice tests in face-to-face courses; our own experience suggests that students could spend as little as 10 minutes or as long as 50 (the entire testing period) for a 30-question exam. It is worth noting that over three semesters the longest any student spent on an exam was only 32 minutes, nowhere close to total time available for the exams.

We evaluated three hypotheses regarding why the students were completing the exams so rapidly compared with our face-to-face courses: first, the exams were too easy; second, the students were exposed to too many questions from the exam pool beforehand; and third, the online environment itself encouraged rapid test taking. We think the first hypothesis is unlikely as the questions originated from a test bank that had already been reviewed as part of the publisher’s quality control process. Both instructors then reviewed the larger initial pool of questions in advance, removing questions that only emphasized memorization of jargon and terminology and selecting questions that were of comparable difficulty and style to those questions used in the equivalent face-to-face courses.

The issue of pool size probably does explain the pattern we see in the exam duration where Exam 1 always took longer than the other exams. The first exam has the largest pool (127 questions), whereas the other exams have comparable-sized pools (98, 82, and 73, respectively) and comparable, shorter, average completion times. We used a “Monte Carlo” computer model to explore the effect of this “pool exposure” on student response times by simulating the quiz and exam question assignments for thousands of trials and determining the average number of previously viewed questions expected for each exam. For example, our model suggests that a student taking Exam 1 will, on average, have seen seven questions previously during the quizzes. To explore the effect this could have on exam duration, we assume that students only “glance” at these questions for a short period of time (5 or 10 seconds) and then recalculate the average time students spend on each new question. For example, for Exam 4 of fall 2011, we calculate a time per question of 21 seconds (total average time divided by 30 questions). From our models, though, we expect that, on average, a student will have seen 12 of the exam questions previously. Assuming they glance at those questions for 5 (or 10) seconds, then the average time spent on each of the remaining 18 new questions increases to 31 (or 28) seconds. Even accounting for the effect of pool exposure, the total time spent per question by our students (column 7 of Table 2) still seems unusually short.

Our final hypothesis is that our students’ apparent haste in completing their exams was related to the environment in which assessment was taking place: online. Ideally one would test this by administering a paper exam to half of the students and asking the other half to take the same test online. Because our course is fully online, we could not do this. However, we suggest that there are a couple of aspects of the online environment that could influence exam times.

First, there has been some research done comparing people’s reading habits online versus on paper, showing that there is a difference in patterns of reading web pages and books (Liu, 2005). This particular study shows that screen-based reading behavior involves mainly scanning and spotting keywords rather than in-depth reading. It is likely that this style of reading also translates to the online testing environment, meaning that our students not only skim previously viewed questions, but also skim all exam questions. Second, there are numerous online resources and books on developing strategies to take multiple-choice tests (e.g., Dobbin, 1984; Landsberger, 2013). These include previewing the test, pacing yourself to leave time to review your answers at the end, critically evaluating all options given for each question, answering easier questions first and spending more time on harder ones, and reviewing all questions and answers at the end. Although these strategies might be used by students taking pen-and-paper exams, our online environment does not seem to nurture appropriate test-taking behavior. Of particular concern is that, given the short amount of time spent by our students on online assessment, it is practically impossible that they backtracked to review answers before they submitted the exam even though this option was available. Exam questions in our course were displayed one at a time, following “best practices” so that students could focus on each question and would be less likely to accidentally miss questions. Unfortunately, the one-at-a-time display does not encourage backtracking because students must navigate through the questions sequentially, and when they have reached Question 30, they might not want to spend the time to review all questions one by one.

Future directions for course design and research

There is evidence that SA is most effective, in that it leads to better
FIGURE 3
Self-assessment (SA) in online modules. Two alternative designs for including SA in an online module are shown schematically. In the sequential model (top), SA is completed as a separate activity done after all of the subject content is covered. In the alternative integrated design (bottom), SA is dispersed throughout the time when subject content is being covered. One possible advantage to the integrated model is that SA can be viewed by the student as another aspect of the content model and not an extra task.

we required it, students focused on getting the points and the right answer, rather than using the tool to learn (Kerton & Cervato, 2012). An alternative design, schematically illustrated in Figure 3 (bottom of figure), is where the SA is embedded throughout the module so that from the students’ perspective, two tasks have been reduced to a single one. An analogy from face-to-face classes is the use of anonymous “clicker” (personal response system) questions to periodically assess student understanding during a lecture. One possible benefit of this integrated design is that students will become comfortable with the idea that reviewing and self-testing are necessary and useful aspects of the learning process. We encourage instructors to try this SA design and encourage LMS developers to create tools that facilitate this approach.

Although probably little can be done to combat the online style of reading, there are ways to improve LMS design for the delivery of exam questions. For example, features allowing students to electronically mark questions they want to review; to have easy access to individual questions, perhaps through the use of thumbnails or a clickable list of question numbers; and to receive system prompts to review questions if time remains would all make online exams more user-friendly.

In closing, we think that the study of student behavior during online evaluation as part of college courses is an area that is ripe for further research. Much of the research on examinee behavior on online exams has been done in the context of large standardized tests such as the ACT, SAT, LSAT, and GRE (e.g., Camara, 2002; Schnipke & Scrams, 2002), and most studies examining student behavior in online courses tend to focus more on issues of cheating and security (e.g., Harmon, Lambrinos, & Buffolino, 2010; Rowe, 2004) rather than on student test-taking strategies and noncheating-related behavior. Because online courses are becoming more prevalent and the LMS automatically and unobtrusively records timing information, more research can and should be done at the individual course level to better understand student behavior in online testing in order to optimize students’ learning experience.

Acknowledgments
We thank Iowa State University (ISU) intern Angela Zhang for her work collecting exam and test timing data. Many thanks also to Joanne Olson from the ISU School of Education for informative and illuminating discussions related to teacher training.

References


Assessment in Online Learning—It’s a Matter of Time


Charles Kerton (kerton@iastate.edu) is an associate professor in the Department of Physics and Astronomy and Cinzia Cervato is a professor in the Department of Geological and Atmospheric Sciences, both at Iowa State University in Ames.

SHARE YOUR IDEAS!

Have an idea for an inspiring presentation or workshop on science education?
Submit a session proposal today for...

NSTA’S 2015 CONFERENCE
to enthuse and stimulate our community of educators!

Chicago, Illinois
March 12–15, 2015

Professional Development Strands:
• Natural Resources, Natural Partnerships
• Teaching Every Child by Embracing Diversity
• The Science of Design: Structure and Function
• Student Learning—How Do We Know What They Know?

To submit a proposal, visit www.nsta.org/conferenceproposals