AC 2010-1150: ONLINE VS. ON-PAPER EXAMS

Edward Gehringer, North Carolina State University
Ed Gehringer, efg@ncsu.edu, is Associate Professor of Computer Science and Computer Engineering at North Carolina State University. His main research area is collaborative learning technology. He received his Ph.D. degree from Purdue University, and taught at Carnegie Mellon University, and Monash University in Australia.

Online vs. On-Paper Exams

Abstract

As information and education continue to migrate to an online format, on-paper examinations are becoming an anachronism. Paper exams ask students to work in an environment that feels artificial—without the information infrastructure that they depend on in their other classwork, and that they expect to use on the job after they graduate. This paper compares a particular form of online exams, the "open-book open-Web" exam with exams taken on paper. The choice of format has a myriad of implications, most of which are not apparent at a glance. This paper considers several categories of differences, including coverage of material, administering the exam, challenges of grading, how to discourage cheating, and the tradeoff between difficulties associated with handwriting and difficulties with coding the exam. It is based on the results of two surveys, one of students who took open-book open-Web exams, and one of instructors who administered online exams.

1. Introduction

In today's world, tests and exams are given in an environment that is increasingly artificial. Most technical work is done with computers. Few people would attempt to write a computer program, analyze forces on a building, or even write a piece of prose, without the aid of a computer. But that's just the kind of environment we place our students in when they take an exam.

All of the leading learning-management systems and textbook publishers have online testing modules that can be used to deliver quizzes or exams. However, few instructors have completely done away with paper exams. Online exams preclude certain types of questions (e.g., recall), while facilitating others (questions based on looking up information and applying it).

There is more to online exams than meets the eye. Usually, some questions are automatically graded. But there are many ways of phrasing a correct answer. It is much more difficult than it would appear to make sure that all legitimate answers are accepted. Usually, manual regarding is needed. In most systems, this requires navigating to every page of every student's exam.

There are major differences in administering an online exam. For example, it's easy to time the test precisely, so that each student gets the same amount of time. But care is required to make sure that some students don't gain access to the answers when they finish early, and then e-mail them to those who are still working on the exam.

This paper covers online exams in general, but focuses on a particular kind of online exam, the "open-book open-Web" (OBOW) exam. This has major implications for academic integrity. An online exam does away with a lot of opportunity for cheating (the kind of cheating that involves consulting unapproved materials or devices), but raises new possibilities, such as electronic communication between students during the exam. In later

sections of this paper, we will consider these issues and more in great detail.

Our data was gathered through two surveys: (1) a survey of the students who took the author's courses in Fall 2009; this survey was administered after the first exam in the course but before the second exam, and again after the final exam at the end of the semester; and (2) a survey of instructors who had used online exams in their courses. This survey was sent to four e-mail lists: the Engineering Technology listserv, etd-l@listproc.tamu.edu, serving ASEE's Engineering Technology division; the SIGCSE members list, SIGCSE-members@LISTSERV.ACM.ORG, serving the Special Interest Group on Computer Science Education of the Association for Computing Machinery; the College Board's AP-CS list, ap-compsci@lyris.collegeboard.com, consisting mainly of high-school teachers of advanced-placement computer science; and the listserv of the Professional & Organization Development Network in Higher Education, POD@listserv.nd.edu, the professional organization for faculty development experts. Approximately 150 responses to the student survey were received, and 85 responses to the instructor survey.

The remainder of this paper is organized as follows. Section 2 is a brief recounting of the author's experiences with online exams. In Section 3, we compare OBOW with paper exams in several respects. Section 3.1 deals with coverage of material. Section 3.2 contrasts the difficulty of reading handwritten paper exams with the challenge of coding online exams. Section 3.3 presents advantages and disadvantages of grading online exams. Section 3.4 covers the contentious issue of academic integrity. Section 3.5 examines the costs and benefits of administering online exams. Section 3.6 surveys several miscellaneous advantages and disadvantages of online exams, such as accessibility and privacy. Section 4 presents numerical results from the student survey dealing with satisfaction, and Section 5 places the OBOW approach in the spectrum of online testing practices. Finally, Section 6 summarizes the results of the paper.

2. A Personal Perspective

Like many faculty, I have migrated my courses to the Web over a period of years. I teach many combined on-campus/distance-education courses, and placing materials online makes them equally accessible to both groups. Students submit most of their homework via an electronic system. Distance-ed students prefer to scan in handwritten homework rather than snailmail it to the DE office. My lecture notes have been online since the mid-1990s. By 2009, the only element of my courses that was not online was the exams. An accreditation visit was coming up, and I knew that online exams would make it easier to tie specific questions to specific learning objectives and show how well the objectives were being met. Inspired by a favorable report [1] from a moderately large study, I decided to experiment with online exams.

I had used the online testing system WebAssign for quizzing in my online Ethics in Computing course for many years. However, the cost structure (a charge of about $20, assessed to each student) made it infeasible to use it for three exams during the semester, especially since one of my courses already had a textbook cost of over $150 for new books. So I decided to use Moodle, the preferred learning-management system (LMS) on my campus. My experience with the first exam was trying. Time expired on students who had not saved their exams. The automatic grading made dozens of scoring errors on at least half

the exams in the class. The testing system "ate" several special characters and truncated students' answers at that point. Fixing these errors took more time than it would have taken to grade the exam manually … and this was after I (and my TAs) had invested more than a day in getting the coding right before the exam. One of my TAs and I created a student survey and submitted it to the campus IRB for approval. It was approved, about 10 days before the second exam in both courses, and administered to the students. The results surprised me. Despite the difficulties, students expressed strong support for the OBOW format. Since the difficulties with the format pertained to artifacts and not to pedagogy, I felt compelled to continue it. This allowed us to gain more experience with coding and administering online exams, and to study how students' perceptions changed as they gained more experience with the format.

3. OBOW Exams: Advantages and Disadvantages

There are several obvious differences between OBOW and on-paper exams. One can't ask simple recall questions on an OBOW exam, but these are frowned on anyway, because memorizing a definition does not constitute very deep learning. OBOW exams offer different kinds of opportunities for cheating. It's much more trouble to code questions for online administration than to use a wordprocessor to format them on paper, but one hopes that this effort will be offset by the convenience of automatically grading objective questions.

These observations just scratch the surface. Based on our experience with online exams and our student and faculty surveys, we were able to identify at least 50 distinct advantages and disadvantages of the two formats. For ease of comprehension, these have been grouped into six distinct areas: material covered, handwriting vs. coding, grading, academic integrity, administration, and miscellaneous.

3.1 Material

Advantages: Because an OBOW exam offers a more real-world environment, it comes much closer to the goal of authentic assessment [2], examining students on the kind of tasks that they will do on the job. It raises the possibility of posing new kinds of questions, such as asking students what specialized resources are necessary to accomplish a job. Providing a list of such resources in a paper exam would waste paper, as well as artificially limiting the number of resources to consider. One might have the students run a particular simulation or animation and explain the observed results. Other kinds of dynamic content can be used, in the words of one of our instructor respondents, "to present concepts and present scenarios to examinees."

The fact that exams are automatically "collected" and frequently automatically graded diminishes the bookkeeping overhead and permits exams to be dispersed throughout the semester. One of our respondents suggested, "Give more quizzes during the semester directly after about 10-15 minutes of content...that is graded by the computer and perhaps students can retake one time." More frequent testing discourages students from cramming for exams, and thus tends to improve retention of concepts.

Disadvantages: OBOW exams essentially prevent recall questions from being asked, because the answer could be obtained by a simple Web search. But the limitation is greater than this. Even explanations can be looked up, so if students are asked to tell why a particular phenomenon occurs, instructors should check that they have not simply cut and pasted their answer. In a large class, this is too time consuming. Even if they have not cut and pasted an answer, they may have reworded one. It is difficult to judge whether students really do understand the concept.

One anecdote is illustrative of the situations that can arise. On one of my final exams in Fall 2009, I noticed that several students gave the same answer verbatim as in my answer key. I had been careful not to place my answer key on any fileserver (since occasionally a student in the class may work for the IT department and have administrative access to my filespace). I suspected that someone had intercepted the e-mail I sent to the TA who was coding the questions for Moodle (I sent her all the answers, both for the automatically and manually graded questions). However, when I confronted the first of the suspected cheaters, he said he had found the answer by a Web search for the exact wording of the question. I tried that, and pulled up a 10-year-old exam of mine, from another course! My policy is to give students access to one old exam to help them study, but the old exam is supposed to be removed from the Web soon after the new exam is given. In the case of the 10-year-old exam, I had obviously forgotten to do that. Since students were explicitly allowed to search the Web, this was not a violation, and I had to give them credit, despite the fact that there was no evidence that they understood the concept.

Online exams almost completely preclude asking questions that require a diagram for an answer. The sole exception would be when all students have tablet computers and the tablet software is interfaced to the testing application. Since diagrams are frequently used to represent processes in engineering, this constitutes a major limitation.

Many multipart questions consist of some parts that have well defined answers and can be automatically graded, and other parts that take freeform answers and must be manually graded. Not all systems can handle this. Moodle, for example, has built in at a low level the notion that a question is either automatically or manually graded. Judging from answers to our instructor survey, it appears that Blackboard (including Blackboard Vista), Desire2Learn, and Sakai do support questions that are part automatically, part manually graded, but some caution must be exercised, because not all respondents agree that the systems can handle this kind of question. For systems that can't handle them, it is not pleasant to design a workaround. Our workaround was to (1) create separate questions for the automatically and manually graded parts, (2) assign our own numbers (such as 2a and 2b) to these questions, and (3) disable the question-scrambling feature of the testing system to assure that the questions would be juxtaposed on everyone's exam.

Summary of material issues	
Pro-online	Con-online
• Closer to authentic assessment • Allows more research & application questions • Allows exams to be more frequent and shorter	• Hard to tell whether students understand an answer or have simply cut & pasted it • Can't ask questions that require a diagram for an answer • Some systems can't handle questions that are part automatically graded and part manually graded

3.2 Dealing with handwriting vs. coding for online administration

Advantages: Online exams free both students and staff from the need to deal with handwriting. One of our instructor respondents noted that it was "[e]asier to type out answers than handwrite them. This is really true if the student has the ability to use a spelling and grammar-checking tool within the test." Another way of putting it is that students can spend more time thinking and less time writing. For faculty, obviously, a major advantage of online exams is the fact that they don't have to decipher students' handwriting. This removes one of the major time sinks of manual grading.

Disadvantages: Offsetting the benefits of not having to deal with handwriting are the costs of coding an exam for the testing engine. If the question requires complicated formatting (involves filling in blanks in a table, for example), it is generally not possible to do this with a WYSIWYG editor. The editor in the testing system often does not support creation of tables, and an HTML editor such as Dreamweaver or FrontPage won't support the commands used for automatic grading. Typically, one needs several attempts to get it right. This is much more demanding than using a WYSIWYG wordprocessor for a paper exam.

Beyond that, the typical online testing systems is unable to recognize minor variations of answers to automatically graded questions. Simple misspellings, pluralization, variations in capitalization, and so forth, lead to answers being marked wrong.   It is difficult for instructors to imagine all the variations of how a word or phrase can be written. In the code, the list of alternatives becomes long, hindering readability. And still some correct answers are marked wrong. Either the instructor needs to examine each answer individually, or (s)he needs to wait until students complain. There is no easy way out.

When answers are longer than a single word, the difficulties compound. Many of my classes involve programming. My students had problems with Moodle, which treats embedded blanks as significant. For example, a blank after a parenthesis in an expression can cause the system to give zero credit for the answer. Moreover, several special characters, including "<" and ">" are discarded by the system[1] before answers are graded. Worse, the appearance of such a character truncates the string at that point. Not only is the student's answer marked incorrect; it is not even saved! Our workaround was to tell students to use "&LT" and "&GT" instead, but not all students remembered to do this at all times.

---

[1] Many Web-based systems remove special characters because it is an easy way to prevent scripting attacks.

This problem is not limited to Moodle. One instructor who used eCollege reported,

> In a web-based environment, students who use double quotation marks within an answer could render a response to a question, or the entire exam, un-readable, until admin staff correct the issue.

If programming classes are hard for online exams, applied math is no easier. A Blackboard Vista user commented,

> One of the serious problems is Equation Editor that cannot be conveniently used in WebCT that I am using.

A Blackboard user sang the same song:

> Applied math questions difficult to implement and for students to use, CMS Java scripts can be slow. Graphing and graph interpretation can be difficult.

And this is true for the students as well as the instructor. An Axio user said,

> Electronic exams are not very conducive to capturing student process on long mathematical problems, due to the difficulty of keying in mathematical sequences.

Screen layout poses additional hurdles. Some systems, such as Axio and Examview, provide text boxes of only one size. This means that if the question is multipart, the beginning of the question tends to scroll off the screen before the last text box can be filled in. (This is also true of the standard release of Moodle, but Moodle is open source, and a small mod to the application can allow textboxes to vary in size.)

Unlike paper exams, it's not possible to see in advance how an online exam will appear to each test-taker. It depends on the resolution of the user's screen, the size of the browser window, and the font size. This causes students to miss questions. As one of our students put it,

> Make questions much more apparent. I actually missed one of the fill-in-the-blanks because it was waaaaaaaaay on the right side of the screen.

One of our instructor respondents suggests that online exams always be pretested:

> The biggest problem was if a question and its answer key were missed keyed on input. We eliminated this by performing pilot testing with one or two students and as well as other faculty to help catch errors before the launching of the exam on line.

Summary of handwriting vs. coding	
Pro-online	Con-online
- Easier for students to type than write - Easier for faculty to read typing than handwriting	- Time-consuming to code questions correctly - Automatic grading of short-answer questions is error prone and must be checked manually - System may discard answers that contain arbitrary special characters - Equation editors are hard for students and instructors to use - Screen layout of questions may be unreadable on some browsers

3.3 Grading

Advantages: We have touched on some of the vagaries of automated grading in the previous section. When it works, it is a godsend, especially in large classes. In fact, this tends to encourage multiple-choice questions in large classes, since autograding of these is more reliable. This may not be the best approach pedagogically, but it is *not* a consequence of online exams; the same tendency occurs with paper-based exams.

But time savings is only one of many grading advantages inherent in online exams. Grading may also be perceived as more objective, because the same automated rules are followed for everyone. Furthermore, feedback can be immediate for automatically graded questions. When it works, this can be a pedagogical advantage. As one of our Blackboard instructors put it,

> Their exams are graded faster. Thus giving them their feedback more quickly enabling them to research their errors while the topic is still fresh in their memories.

On exams that are partially manually graded, though, immediate feedback can give a misleading impression of a student's score. Before manual grading takes place, what grade should the system present to the student? Some systems give 0 credit for questions that have not been graded; others give full credit. Either one can be far off the mark. Sometimes it is easier to disable immediate grade feedback than to deal later with dashed expectations.

In any case, though the student gets the feedback as soon as the instructor releases it. This contrasts with paper exams, where the student has to wait until the next class period to see the instructor's comments. Moreover, it allows students to look over their final exams for grading errors. With paper exams, students often leave town before they have this opportunity. This can be a two-edged sword, however, because it inhibits faculty who have been using the same final exam year after year.

Many systems not only give students their own score, but also show them how they compare with others in their class. This can be useful information, which helps allay concerns when an exam has been unusually difficult.

When several students make the same set of errors, instructors get tired of writing the same comments on each test paper. With an online system, they can save a set of standard comments, then cut and paste them into various students' exams, as appropriate. This saves time and encourages copious feedback.

Online exams facilitate grading in other ways. It allows instructors to grade a question at a time, on all exams, before moving on to the next question on any exam. This makes it easier to construct and remember a mental rubric, and thereby promotes fairness and repeatability in grading. It may also save time, because the instructor does not lose context in switching from question to question.

In a large class, online exams make it easier for the instructor and TA to grade in parallel. They do not have to pass papers back and forth. They don't need to worry about who is going to take the papers home over the weekend. If an instructor finds that a question has been missed on a paper, (s)he can quickly assign it to the TA who missed it without

physically delivering the paper. This is a big advantage when the grade-submission deadline is impending!

Online systems also take care of automatically recording a student's grade. It's not necessary to shuffle through the papers to record the grades in a gradebook. Transcription errors are precluded. One does not have to worry about a TA returning the papers to the students without recording the grades!

The last advantage of online exams is that they provide more data for assessment. It is easy to check how students have done on particular questions that relate to particular learning objectives. This helps at accreditation time. And access to data helps an instructor to see how a class is doing. One Vista user praised the system for "provid[ing] powerful analysis tools." Instructors can use this information to improve their questions over time:

> Another benefit is that individual questions and their answer responses can be evaluated over time. If a question consistently pulls an incorrect response over a period of time it might indicate that question needs rewording or its answers may need better clarification. It could also mean that maybe that topic is not being adequately covered by the faculty, which is causing student responses that do match the correct response.

Disadvantages: As we have seen, automatic grading is frequently a mirage, because short-answer questions often need to be examined individually. Another Moodle user put it this way: "larger effort required to do manual grading than paper tests."

Some instructors actually change the question format to facilitate accurate automatic grading. A Desire2Learn user explains it like this:

> It is harder to develop good questions that don't require re-grading what the computer grades as being incorrect. You can develop good matching questions that make the student work through a problem to get to an answer but it is harder.

Section 3.2 recounted some of the difficulties of coding mathematical questions. They are also harder to grade because, as one instructor put it, "for involved problems you don't see the student work (I don't use electronic exams for these types of questions)." A user of the OnCourse LMS said that at his institution, to get partial credit, students need to scan in or fax in calculations to receive partial credit.

The lack of partial credit may impact class averages. A Blackboard Vista user said, "usually grades are worse … stiff penalties for simple math errors, no partial credit."

Not all testing engines handle rounding correctly all the time. A Blackboard user complained, "Some algorithmically generated numbers made the rounding instructions very difficult, ie [sic] requesting two decimal places when with these numbers, the answer terminates with one decimal place."

Testing engines integrated with LMSs can populate the LMSs gradebook automatically. But if your gradebook is a different piece of software than your testing system, you must move the grades from one to the other. I tried to devise an automatic procedure for doing so, but this was not straightforward, because Moodle exported the grades by name, whereas the gradebook imported them by student ID. This led to problems in corner cases, such as

Summary of grading issues	
Pro-online	Con-online
- Automatic grading may save time - Automatic grading perceived as more objective - Immediate feedback on automatically graded questions - Students have more of an opportunity to look over graded final exams - Students can quickly see how they compare with peers - Instructors can more easily make "standard" comments on answers - Grading a single question at a time may be more efficient - Easier for multiple graders to interact - Grades are automatically recorded - More data is available for assessment and accreditation	- Automatic grading may waste time - Harder to give partial credit, especially on mathematical questions. - Rounding may be tricky to handle - If an external gradebook is used, export/import of grades may be tricky

two-word last names, and resulted in some exam scores not appearing in the gradebook.

3.4 Academic integrity

Perhaps no single issue received as much attention in the instructor survey as academic integrity. OBOW exams remove some modes of cheating, while opening others. It is not possible to cheat by using unauthorized materials, because all materials are authorized. One of my students described it this way:

> This exam does cut down on cheating because if 1) Student gets outside help, well they have their book/notes anyways, the only thing they are gaining is time, which arguably is a non-issue. or 2) Student gives prior notice to friend about questions, still the only thing gained is time, which is still arguably a non-issue.

However, the possibility of friends helping friends must be taken seriously, especially when one friend is a much better student than the other. It's not hard to cheat in this way. One of our instructor respondents related:

> During a University-sponsored workshop for instructors on how to create and deploy on-line testing, I created a sample test and set it up for automatic grading and showing students the correct answers.

> Then prior to "taking" the sample test, I opened up an e-mail application, a couple of extra browser windows, an ASCII text editor, and a simple graphics program. The ease of cheating became apparent when during the test I used the [Alt]-[Tab] application switching function of Windows to activate another application and was able to snap-shot the supposedly "uncopyable" test using the Print Screen key saving the files to a USB drive as I went along. Then, within one (1) minute of completing the exam and viewing the graded test with complete answers (and copying that), I had attached the various screen shots and the answer key to an E-mail and sent it.

> When I explained to the class what I had just done and how easily & quickly I had done it, the workshop leader immediately countered that "our students are just not that technically sophisticated."

> When pressed by the other participants about the limits of on-line test security, the instructor refused to address the issue further. Ignorance is bliss, I guess.

Unfortunately, it is not technologically possible to prevent students from using e-mail or IM during an exam. If the exam were held in a lab, one *could* perhaps block all known e-mail and chat systems on lab computers. But there is no way to be sure that students would not use some channel the blocking software was not aware of. They might even run their *own* chat application during the exam. To lock them out totally would mean that outgoing traffic from their browser would be blocked—even preventing them from typing in search terms, for example, which would defeat the purpose of an open-Web exam.

But the era of lab computers is in its waning years. With laptop ownership near universal among students, colleges and universities are phasing out labs [3]. No longer will the school or the instructor be able to control how students use the Internet while in the classroom. There seems to be little hope of preventing communication during online exams by technological means.

While we cannot prevent students from e-mailing or IMing answers to one another, we can make it harder by randomizing the order of questions on the exam, so that one student cannot ask another for the answer to Question 3, for example. Almost all online testing systems support randomization of questions. But our workaround for not being able to combine automatically graded and manually graded parts in a single question (see Section 3.1) makes it impossible to use this randomization.

Furthermore, one of our Blackboard users said that sometimes randomizing questions *and* multiple-choice answers resulted in the wrong answer ending up with a particular question:

> The system uses a separate program (Respondus) to scramble questions. Blackboard then scrambles answers. Sometimes the two would not interface correctly and the wrong answer was assigned to a question.

Another way of inhibiting cheating is by using different data sets for different students:

> Exams can be repeated. The questions I use require a numerical answer, but each question has 50 sets of data so the student is unlikely to get the same test data. Students can take the exams whenever they want.

But if students can choose when and where to take the exam, how does the instructor know who really submits it? One instructor opined,

> The huge downside to an online format is that it is incredibly easy to cheat. Whether that be talking online to someone or paying someone online to do it for you, or not even being in the room while it's happening, the format is not a good indicator of the knowledge received in the class. A suggestion is to figure out some way to make sure people are only on the test themselves.

Another instructor said that cheating was rampant, in his experience:

> I have been on both sides of the fence on ON-LINE exams. I will tell you that 25-50% of the students cheat [sic] somehow on ALL on-line exams. If there are no barriers possibly more than 50%. OF this I'm sure. I don't give on-line exams at a distance.

One defense is simply to proctor the exam. When I began giving online exams, I decided that all students should be in the classroom, or a designated computer lab, at class time to take the exam. One of my students had this observation about cheating:

> The exam was considered open resource (web, book, note, java api etc.). I'm assuming that by cheating you mean students could conference using web chat programs and such. I took the test in the reserved lab and the lab assistant constantly checked on us for cheating however I'm not sure what precautions were taken in the actual classroom for students who brought their laptops.

Another instructor advised, "Sit behind the students. I sit behind them. They can't see me so they don't know if I'm watching." Software (SynchronEyes, LANSchool, or ABTutor Control, etc.) is available for labs to allow an instructor to pull up any student's computer screen on his/her monitor. It's designed to make sure that students are following the class material rather than e-mailing or Facebooking during class, but it can just as easily be used to proctor exams. But it would be an administrative nightmare, at best, in an environment where students took the exam on their own laptops.

Most systems allow the instructor to configure the exam so that students can't go back to questions that they have already answered. Clearly, this increases the difficulty of cheating. But "some students have complained about not getting to go back and answer a previous question based on what they may have learned in a later question," reports an instructor who uses this approach.

Security can be improved if one is prepared to forgo the OBOW environment. Software such as SecurExam allows computers to be configured so that they can only access the exam, or only access the exam and a list of other pages that are specifically whitelisted. One instructor told us, "you can check IP addresses and time test was taken. Our lockdown browser prevents students from IM-ing or using any other item on the computer except the test (can't print, e-mail, etc)." Finally, to reduce the motivation to cheat, some instructors simply use paper for any test that is worth more than a few points of the student's grade.

A related issue arises when students save a copy of the exam. This does not gain them an advantage, but it can be used to help build a "test bank" for their fraternity or other group. Of course, test banks have been assembled for generations with copies of paper exams; online copying just makes it more convenient. Nonetheless, the effect is to increase the burden on the faculty to come up with more new questions. An instructor at a large public university in the West told us,

> Without a large test data bank students team and eventually have a printed copy of all exam questions. Again, good for covering material but not really good for an individual assessment.

In summary, there is wide concern over the integrity of online exams. Problems can be mitigated, though not eliminated, by proctoring the exam. But as will be seen in Section 4, most students do not believe that OBOW exams are more susceptible to cheating than paper exams.

Summary of academic-integrity issues	
Pro-online	Con-online
- Can't cheat by using unauthorized materials - Can randomize questions and answers to inhibit cheating - Can use different data sets to inhibit cheating - To discourage communication, exams can be proctored - To inhibit cheating, can prevent students from going back to earlier questions - Browsers can be "locked down" to prevent communication	- Easier to cheat by communicating with others - Randomization does not work well on all systems - Easier for students to get copies of all questions - Proctoring not possible for students who take exam at a time or place of their choosing - Students find this constraining - Locking down browsers defeats the purpose of an open-Web exam

3.5 Administration of the exam

Advantages: Online exams can afford all students the same amount of time. No one gets the test paper before another; no one can try the instructor's patience by staying late.

Online exams allow flexibility in timing; students can be allowed to take the exam at a time of their choosing:

> It allows students to take an exam when they have the time or if they cannot make it to class. Some students are not "fresh" after working an entire day and may not do as well if they took an exam later in the evening.

They can save class time:

> Instructors don't have to "waste" large blocks of precious class time on periodic evaluation exercises; more classroom time for lectures, discussions, and labs.

Freed from the constraints of a class period, instructors can give students almost an unlimited amount of time to work on the exam. But of course, all time-shifting comes at a cost. It makes it difficult or impossible to proctor the exam, and thus facilitates cheating.

Online exams allow location-shifting too. Students can be allowed to take exams off site while on business travel or deployed in the military. I had one student take an exam while in India because of a family emergency. While location-shifting also raises the possibility of cheating, if only a few students are permitted to take the exam off site, the risk is not too great. Special attention can be given to these students' exams; their scores can be compared with other exams that they take face-to-face, and they are aware that they can be singled out for attention.

Disadvantages: Online delivery makes it more difficult to give an exam to a student who is not registered for the class section. An instructor may need to do so when a student is finishing up an incomplete, or taking an exam at a different time or with a different class

section because of a schedule conflict. It's not possible to simply hand the student the exam paper. A separate exam may have to be created in the testing system for just this student.

If the student is taking the exam at a different time, one must be careful to set it up so that only that student can access it, and so that the answer key is withheld until all other students have taken the exam. The easiest way to do this is to put a password on the exam and only tell the password to the student authorized to take the exam early. But, how does the instructor know that this student won't e-mail the password to a friend, giving the friend a sneak preview of the exam? The instructor can probably discover this by checking the log, but how many instructors will bother? Or, what if the instructor forgets to change the password after the early student finishes?

Exams do not work quite the same way in all browsers. In the Moodle quizzing system, we found that text boxes did not appear in the same place on all browsers. On some browsers, they covered up parts of the question. Students using Google Chrome had other problems. The exam window tended to freeze up, especially when scrolled horizontally. These students needed to be given an opportunity to retake the exam. On later exams, we warned students not to use Google Chrome.

Two students described their difficulties:

> The online format's all right... the problems are just inherent to using forms like this one. I had some bad experiences with window closings, the refresh key, and a button on the sides of the mice that's apparently been pre-programmed to act as the back button. I'd bet those blanked pages frightened a lot of people. I wish I had an idea how to make those relatively harmless, but other than that, the procedure worked well.

> Most of my complaints are around the Moodlepilot system. I found myself starting with chrome and being unable to scroll. I then switched to firefox which seemed to work well, however, I found myself using the arrow keys to scroll and at times, this would cause my radio buttons in the answer to change instead of scrolling, so a lot of time was spent verifying that I didn't change the answer by mistake.

At my university, the distance-education program approves proctors for off-site exams. Normally the proctor is another employee of the company that the student works for. Online exams pose problems for students who work at security facilities:

> Another problem I encountered is that I work in a secure environment and I have my exams proctored by our training officer at work (while taking leave). In the secure environment, we are not allowed to install anything on our computers, so this makes certain tasks (such as configuring to be able to access the virtual computing lab, or installing programming languages) infeasible.

At other universities, network connections have been a source of difficulty.

> The biggest problem I have had is students losing their Internet connection part way through an exam. This appears to be the result of spending too much time working a problem by hand and then "timing out."

Another instructor related that when a Web connection was dropped, a student was locked out of an unfinished exam, and had to retake the entire exam after contacting his instructor.

Some of the difficulties attributed to the network may in fact be the students' fault, one respondent said:

> A few have trouble with the network because they don't maintain their computers properly, or they are just not proficient, which are issues that ought to be addressed separately.

Instructors can try to make up for network problems, but how do they know that students are not just making excuses?

> Stupid or lazy students will cheat or find a way to make it seem like the software isn't working or they can't access it. 90% of my students had no problems using the system. Most of those who said they had problems didn't bother to document it. I asked them to send me an e-mail if they had a system problem. Most never did, but they would come to class and whine about it. If I pressed them, they would admit they didn't think about sending an e-mail.

Then, there are problems related to saving work. In Moodle, when the time for the exam expires, the last saved copy of the student's work is submitted, and any unsaved changes are lost. This surprised several students on our first exam. In Blackboard, I'm told, the policy is different: Students who click on "Save" instead of "Submit" at the end lose their work.

In both Moodle and Desire2Learn, the Save and Submit buttons are near each other, raising the possibility of accidentally terminating the exam:

> I was afraid of accidentally exiting the exam. Maybe it is just me :) the save only and save and exit buttons were close to each other and were similar.

If this happens, the instructor can simply print out the exam and have the student answer the rest of it on paper.

Summary of administration issues	
Pro-online	Con-online
- All students have the same amount of time - Students can take the exam at different times - Students can take the exam in different places	- Time-shifting and location-shifting facilitates cheating - It is more difficult to give an exam to a student not registered for the class, or one who needs to take the exam early - Some browsers may have trouble with the exam, or with certain questions - Students may fail to save their work, or accidentally exit the exam - Students in high-security environments may not be able to access the exam - Network problems may abort exam attempts - Students may use imaginary network problems as an excuse for their own lack of preparation

3.6 Miscellaneous issues

Advantages: Some students find using a computer to be less stressful.

Another issue is privacy. Since exams are not put in a pile to return, there is less chance that one student will see another student's grade.

Online exams make it easier to reuse questions. This is true for two reasons. Most systems provide a question bank that allows the same question to be included in multiple exams. Also, questions can easily be "mixed and matched," so even if students have access to old exams, it will be difficult for them to locate the questions in time. If there are enough questions, questions can be scrambled so that no two students receive the same exam.

Finally, online exams have been called "more environmentally friendly." In terms of cost, though, savings in paper have to be offset by increased IT support costs.

Disadvantages: Some students find using a computer to be *more* stressful:

> Some students, even in tech classes, find it hard to struggle with the environment at the same time they are being tested on the material. Some are flustered with electronics and feel pressured. Some ADHD students really have a hard time with the screens, I found screen layout had a big impact on them. They had a hard time going back and forth and losing focus.

The ADHD allusion raises the issue of accessibility. Students with certain repetitive-strain injuries will also be disadvantaged if they have to take exams on a computer. On the other hand, blind students can use screenreaders, which is less trouble than accommodations that would need to be made for paper exams.

Online exams may also pose problems for students without disabilities:

> I prefer a paper format. For me:
>
> 1) with a paper test, I can easily scan over the test to see what it contains. I can better plan my time, and concentrate on the sections that are the most important.
>
> 2) with a paper test, I can doodle and draw next to the questions to help me understand them better. I guess I could have written on separate paper, but tying the paper with the questions number was less attractive.

Open-Web exams may lead students to spend too much time browsing for an answer, and too little time writing up their answer. So many students told us this, that we began subsequent exams with a warning against browsing too long. This finding is consistent with the observations of Boniface [4] and Ioannidou [5] (as quoted by Rakes [6]), who found that some students performed more poorly on open-book than on closed-book tests, because they spent time looking through their textbook or notes. Rakes found that instruction in how to prepare for open-book tests improved students' performance. This suggests investing more effort in studying how to prepare for online examinations.

Offsetting the privacy advantage is the fact that if the instructor never has to return homework or exams to students, (s)he loses a valuable opportunity to learn the students' names.

Finally, with your questions in an online system, they are stored in a centralized database rather than being on your hard drive or fileserver. This means that if your institution changes switches to a different system, someone has to port your questions, or they are lost to you. Ease of porting is not guaranteed. In the worst case, you might have to bring up your questions one by one on the old system, and copy and paste them into the new system.

Summary of miscellaneous issues	
Pro-online	Con-online
- Some students find using a computer to be less stressful - Privacy—students have less chance to see others' graded exams - Easier to reuse questions - More "environmentally friendly"	- Some students find using a computer to be more stressful - Accessibility—students with ADHD or RSI may be disadvantaged - Harder to get to know students' names without passing back exams - Harder to scan through test at a glance - Students may waste time browsing Web for an answer - If institution switches to new system, may be difficult to retrieve and use old questions

4. Results of student survey

Results of the student survey are given in the table below. There were 115 students in the undergraduate class (two on-campus sections) and 84 students in the graduate class (on-campus and distance-ed sections), so the response rates were about 50% for the first survey and 30% for the end-of-semester survey. All questions were rated on a Likert scale from 1 to 5, with 5 signifying "strongly agree." Both classes thought the questions were relevant to the material covered in class, though the undergrads' agreement tended to tail off at the end of the semester. Students were uniformly positive in reaction to the OBOW format, giving it average scores between 4.12 and 4.5. The graduate students thought it was more beneficial to be able to execute code than the undergrads did. They also thought it was a more authentic testing environment (Question 4). The graduate class apparently believed that the later exams went better than the first exam, while the undergraduate class thought the opposite. Neither class was particularly concerned about cheating, though both classes were more worried about it after they gained more experience with the format.

	Undergraduate class (CSC `216)			Graduate class (CSC/ECE 517)		
	Avg.	After 1st exam	After final exam	Avg.	After 1st exam	After final exam
Number responding	79	50	29	73	53	20
1. The questions were relevant to the material covered in class.	3.86	4.02	3.57	4.22	4.12	4.5
2. The test format of open-book, open-web was beneficial.	4.29	4.32	4.21	4.44	4.38	4.6
3. The ability to execute code was beneficial.	3.61	3.59	3.64	4.44	4.38	4.6
4. The format was relevant to the way work is done in business/profession.	3.75	3.92	3.46	4.19	4.08	4.5
5. Taking the test online worked well	3.81	4.02	3.42	3.53	3.17	4.5
6, The exam structure allowed students to cheat.	2.14	1.90	2.57	1.76	1.69	1.95

These findings can be compared with those of Williams and Wong [1]. Their population was different from ours, in that their students were enrolled in a program where all courses ("subjects") used OBOW exams. Most of them were between 30 and 39 years old, considerably older than even our graduate students. Their OBOW exams were flexible in time and place, unlike our exams, which were proctored at a particular location at a single time (DE students could take the exam at any time during the day, but were required to take it on the same day as the on-campus students).

Our survey included more students than theirs (n = 54 for them; we had 152 responses from at least 103 distinct individuals). It had three questions in common with theirs. Their students rated Question 1 ("aligned with pedagogy" in their terminology) an average of 4.34 compared with our classes' 4.22 and 3.86. Their higher agreement may be reflective of the questions on their exams, which were related to a particular case study. On Question 4 (relevancy of format), their students rated OBOW exams 4.42, compared to our students' 4.19 and 3.75. However, they were more concerned about cheating on the exam (2.83) than our students (2.14 and 1.76). This may be related to the fact that they had complete freedom about when to take the exam, especially that students consider take-home exams [7] more susceptible to cheating.

Thus, ours is the second study to show high student satisfaction with the OBOW format.

5. Other Online Testing Formats

Several universities have experimented with a combination of paper and online exams. A well known computer-science educator explains the rationale by saying that online exams can test students' ability to write code, while paper exams can ask them conceptual questions, without getting them bogged down in the details of programming-language syntax. Of course, programming is a special case, because having students write programs during an exam is about as clear a case of authentic assessment as one can imagine. A study by Woit and Mason [8] shows that weekly online programming quizzes are more effective in motivating students to perform well than regimes in which work is assessed less regularly. Bennedsen and Caspersen [9] also report favorable results with online assessment in programming courses.

Online exams that are not OBOW appear to be fairly common. Just about as many of our 85 instructor respondents said they had used the OBOW format as had not. Ten respondents reported using software that prevented the students from accessing the Web during the exam.

Another possibility is to give a paper-based exam to students who are in front of a computer that can browse the Web. This avoids all of the administrative and coding problems inherent in the OBOW format, as well as most of the material disadvantages (one can easily ask objective/subjective multipart questions, or questions that require a diagram for an answer), while still preserving many of the benefits of an OBOW exam.

6. Summary

Online exams offer a promising environment for assessing students' knowledge and skills. OBOW exams, in particular, can provide a more authentic assessment environment. At first glance, one might assume that the largest advantage of OBOW exams is time savings, but this is not correct. In most of today's systems, it may actually take longer to give an online exam than a paper exam, especially when coding time is considered. Online exams raise a large number of administrative problems that have no parallel in paper exams. However, there is strong evidence that students prefer OBOW exams. Online exams are better for getting feedback to students in a personalized and timely manner. Since many of the advantages of OBOW exams are inherent, while many of the disadvantages are artifacts of current technology, we can expect OBOW to play a larger role in curricula of the future.

Bibliography

[1] Williams, J.B. and Wong, A. 2007. "Closed book, invigilated exams versus open book, open web exams: An empirical analysis" *Proc. ASCILite 2007*, pp. 1079–1083.

[2] Wiggins, G. 1990. *The Case for Authentic Assessment*, ERIC #ED328611.

[3] Terris, B. "Rebooted computer labs offer savings for campuses and ambience for students," *Chronicle of Higher Education*, December 6, 2009, http://chronicle.com/article/Computer-Labs-Get-Rebooted-/49323/

[4] Boniface, D. 1985. "Candidates' use of notes and textbooks during an open-book examination," *Educational Research* 27:3, pp. 201–209.

[5] Ioannidou, M. K. 1997 "Testing and life-long learning: Open-book and closed-book examination in a university course," *Studies in Educational Evaluation* 23:2, pp. 131–139.

[6] Rakes, G. C. 2008. "Open-book testing in online learning environments," *J. Interactive Online Learning* 7:1, pp. 1–9.

[7] Weber, L. J.; McBee, J. K. and Krebs, J. E., "Take-home tests: An experimental study," *Research in Higher Education* 18:4 (December 1983), pp. 473-483.

[8] Woit, D. and Mason, D. "Effectiveness of online assessment." *Proc. SIGCSE 2003: 34th Annual Technical Symposium on Computer Science Education.*

[9] Bennedsen, J. and Caspersen, M.E. "Assessing process and product—a practical lab exam for an introductory programming course," *ITALICS* (*Innovation in Teaching and Learning in Information and Computer Science)* 6:4 (October 2007), pp. 183–202.