

Assessment in online courses: Some questions and a novel technique

Simon

University of Newcastle, Australia
simon@newcastle.edu.au

***Abstract:** When students have little to lose and a great deal to gain by cheating, some of them will do so. Students have even less to lose if caught cheating in remote online courses than in face-to-face courses, and so are more likely to cheat in online courses. This paper describes an electronic 'watermarking' system that we have used to detect a particular form of cheating in remote online exams. Findings from the use of the system raise some disturbing questions about the practice of distance education and education in general.*

***Keywords:** assessment, authentication, distance learning*

Introduction

It is customary in the literature of higher education to adopt the polite convention that everybody is honest. Of course there is the occasional acknowledgement that students will collaborate on individual assignments or copy and paste essays from the Web, but in general, we seem to accept that the work submitted by a student is the work of that student. When authors do acknowledge the possibility of cheating, they sometimes suggest that it might be the teachers' fault. Sheard, Carbone, and Dick (2002) found that 80% of students admitted to some form of cheating. They concluded that students should be assisted to 'develop strategies to manage the internal factors that lead to poor learning tendencies', and that 'it is also important for educators to address external factors . . . caused by characteristics of the learning environment they provide for students.' What appears to be missing from such studies is a clear acknowledgement that when students have little to lose and much to gain by cheating, some of them will invariably choose to do so. There is no magic wand that will make all students choose honesty over dishonesty regardless of the outcome.

The little that students stand to lose by cheating appears to be further diminished by remoteness. When they never see their teachers or administrators, when everyone else involved is reduced to text messages at the other end of an Internet connection, there appears to be less concern about loss of face. With the typical remote program being a short postgraduate offering, the average student has not invested a great deal of time and money in the program and can readily switch to a different program at a different institution if expelled. In our web-based remote courses, we therefore chose to ignore the polite convention that everybody is honest, and set out to see if we could detect blatant cheating. We created a system to detect certain forms of cheating on our remote online exams. We were disappointed, but not surprised, when we succeeded.

The purpose of exams

In many University courses, the final examination is the ultimate assessment of the student's knowledge and skills pertaining to the subject matter. It is the means for students to show what they have learnt, for academics to assess what the students have learnt (Thorpe, 1998), and perhaps even for some cynical academics to show the students how little they have learnt. But many computing courses have a strongly practical orientation, which is better assessed in assignments (formative assessment) than in exams (summative assessment). Why conduct exams in these courses?

A common answer (eg Thomas, 2001) is that while we can never be sure just who has written a student's assignments, we can almost always be confident that a student has written his or her own exams. For this reason, even in highly practical computing courses such as programming or database implementation, we tend to set exams and to allocate a high proportion of the final mark to them. This confidence is occasionally misplaced, and one hears now and then of students paying other people to take their exams (Zobel, 2004; Fröhlich, 2000); but the general principle still applies, and we set exams to improve the likelihood that we are assessing the work of the right person.

The value of this approach is still under discussion. Some authors suggest that 'current assessment practices in higher education are long overdue for a re-think; they are particularly ill-suited to the digital age' (Mason, 1989, as cited in Ruhe, 2002) and that success depends on 'a progressive shift from summative to formative approaches and reaching a balance of both' (Lim et al. 2004). The opposing viewpoint is that 'a distance learning programme must be based on an existing [face-to-face] programme, which is the gold standard for curriculum and its assessment, and where possible, common examinations and assessments [should] be used.' (Clarke et al. 2004).

Notwithstanding the debate, many academics continue to use exams in the belief that they are more readily authenticated than assignments.

Techniques for conducting exams in remote online courses

Thomas (2002) observes that 'a common criticism of remote examinations used for summative purposes is the difficulty of ensuring that cheating is minimised'. Universities have developed a number of techniques for conducting distance exams with this in mind. Some of them are surveyed briefly in this section.

The cooperating institution

It is not uncommon for a University to ask cooperating institutions in the students' neighbourhoods to invigilate exams. This technique offers about the same level of security as an exam invigilated at the University itself. Unfortunately, it takes time and effort to establish a network of cooperating institutions. Thomas (1998) observes the difficulty faced by the Open University in making special arrangements for students outside the UK.

The advent of Web-based online courses has spread the student body so wide, and yet so sparse, that the establishment and maintenance of a usable network of institutions would be all but impossible. By way of illustration, in one of our recent courses it would have been easy to arrange invigilation for the 15 students in five different locations within Australia, and reasonably easy for the dozen in Singapore and the half dozen in Hong Kong. But it would have been somewhat more difficult to find cooperating institutions for the students in Botswana, Central Africa, China, Malaysia, Netherlands, Norway, Papua Niugini, and the United Arab Emirates.

The trustworthy third party

A number of Universities ask the students to arrange a trustworthy third party to invigilate the exam. In one arrangement used at our University, the invigilators have to sign statutory declarations attesting that the student sat the exam honestly. One obvious flaw with this system is that the statutory declaration might easily be completed by an untrustworthy person who has no qualms about lying. Indeed, there is no way of ensuring that the declaration is completed by a third party and not by the student.

The webcam

A researcher at the University of Newcastle, Australia, has developed a system whereby the invigilation is performed remotely with the help of a webcam. No reference is available for this; we believe it is the subject of a patent application rather than an academic paper. While students are doing the exam, the webcam is turned on and transmitting pictures of them to an invigilator back at the University. One simple way to defeat this system is for the student to write a note and move it to one side, out of the camera's field of vision, where it is picked up, answered, and returned to the student. Another is to email for help, unless the webcam software also transmits screen images of the student's computer to the university. Fröhlich (2000) reports on a number of technological systems, including a patented 360° webcam system and other options such as biometric recognition, fingerprint recognition, retinal scans, and others. There is potential value in all of these, but they can all be subverted in various ways.

Trust

Some Universities trust the students to do the exam honestly and fairly. That is their choice, but it does rely on the polite convention mentioned in the introduction.

Constraints on exams in remote online courses

In traditional university courses an examination is held at a specified time and place, and all students attend. Alternative arrangements might be made for students prevented from attending by dramatic and unforeseen circumstances, and even sometimes for students prevented from attending by foreseen circumstances, but for the bulk of the students the exam time and place are fixed.

One of the advantages of an online course is that students can 'attend' at times that suit them. Indeed, it is for this reason that students will sometimes prefer an online course to a face-to-face one offered by an institution in the same city they live in. It is a reasonable expectation that this flexibility of attendance will extend to an examination if there is one. Further, some of the students will be studying from home while others study at work. Further still, in a given online course the students will be spread across different time zones. The class mentioned briefly in the previous section took in six different time zones with a span of 10 hours. For these reasons it is clearly impractical to conduct the exam at a single time (what Thomas (2002) calls a 'synchronous examination'), and we must resort to an 'asynchronous examination', in which students choose their own sitting time from a specified range.

Conducting a timed remote online exam

We have recently run asynchronous timed exams in the following manner. The exam is available for 48 hours, covering the best part of a Sunday and a Monday in most time zones, to cater for most student timing constraints. Students are asked to download the exam some time during the 48 hours, to complete it by writing their answers in the same document that was downloaded, and to return it. They are given a time limit that includes the notional exam time and an additional 15 minutes or so to allow for internet connections, downloading, and uploading. Exceeding the time limit is penalised, typically at 2% for every minute by which the limit is exceeded.

When students are ready to download the exam they direct a browser to a specified URL, where they find a password-protected form. When they click a button on this form, an email is sent automatically to the lecturer, with the time, the student name, and the student number. The student is then taken to a second form containing a link from which the exam document can be downloaded. The student adds answers into the exam document, then returns it via the Digital Dropbox feature of Blackboard™, which puts a time stamp on its return. The download time, from the automatic email, and the return time, from Blackboard™, are combined to establish the time taken by the student to complete the exam.

Some ways of cheating in a remote online exam

As in most academic situations, the number of ways of cheating is limited only by the imaginations of the students. Students might try to bypass the form that sends the automatic email, thus accessing the exam without the lecturer's knowledge. Having prepared their answers at their leisure, they will then download the exam in the expected way, complete it, and return it in good time. A simpler way of arranging early access to the exam is to have it sent on by a fellow student who downloaded it early in the period of availability. Students might cheat by having other students or mentors with them when they do the exam, and taking advice or assistance from these extraneous people. Students might cheat by emailing other students for assistance during the time they are doing the exam.

The electronic watermark to detect one form of cheating

Believing that early access to the exam was perhaps the likeliest form of cheating, we devised a system that might help to detect such cheating. The principle of the system is that there are many versions of the exam paper, indistinguishable to the students, and that different versions are made available at intervals during the period of availability. When a student's exam is returned, it is checked to see whether it is the same version that the student officially downloaded. If it is an earlier version, the student must have had access to the exam before downloading it.

The distinction between versions, which we call an electronic watermark, consists of a number of elements. The simplest is a double space in a different position on each version of the paper. Somewhat trickier is a version number, invisible by virtue of being white on white, in a part of the footer where very few people would find it. Both of these clearly mark each version of the question paper, but neither marks the text of the answers. If a prudent student were to prepare answers in an early version of the paper, then copy and paste them into the version downloaded, these measures would not detect this.

To address this problem, the principal element of the watermark is the text colour. All the text of each version is in a colour that is visually indistinguishable from black, but distinct from every other version. When students place the cursor into the paper and type an answer, that answer acquires the same colour as the questions. If a student were to copy and paste answers from an early version into the legitimate version, the questions would be in the colour of the legitimate version but the answers would be in the colour of the earlier version. The system has been described in more detail in a more technically oriented paper (Simon, 2005).

Successes of the watermark system

We have now used this system for three exams. As well as detecting the expected forms of cheating, the system has helped us to uncover a ring of cheats in which several students worked together to answer the exam for one student, then the others took it away and submitted their own answers up to 25 hours later.

We have also detected a repeating student who copied one of his answers from the previous year's paper, and a student who copied one of his answers from the paper of a friend who did the course the previous year. Because the exam is open book (for obvious reasons), it is not clear that this constitutes cheating.

Questions arising from use of the watermark system

Is it worth the effort?

The system has clearly shown that a number of students are willing to cheat on their exams, and has helped us to deal appropriately with some of those. But the cost is significant. In its current form, the watermarking system requires almost constant email vigilance over the 48 hours that the exam is available, as we have to upload a new version of the exam every time we see that the previous version has been downloaded. This has the effect of leaving the lecturer sleep-deprived, which is not the ideal state in which to begin marking exams.

The system succeeded almost spectacularly in its first use, and has not definitively detected any cheating since then. Has word passed around that we are indeed able to detect cheating, and has the cheating subsequently stopped? We believe it more likely that students have some idea of what sort of cheating we can detect, and that they are managing to cheat in other ways. If this is the case, we are going to all the effort of creating multiple versions of the exams and 'baby-sitting' the email for 48 hours with little likelihood of a useful return.

It would clearly help if we could automate the delivery of distinct versions, a goal that has so far eluded us. But even if we eventually succeed, there will still be the tedium of preparing and double-checking as many versions of the paper as are required.

Is there any point in conducting exams in remote online courses?

It was argued earlier that the principal reason for setting exams as well as assignments is to acquire some work that we can be sure is the student's own, produced in conditions that are more or less comparable with those in which other students are working.

It is evident to us that students are willing to cheat in online exams. We set up a system to detect one form of cheating, and it succeeded immediately. We assume that other forms of cheating, not detectable by this system, are also taking place. We are left to conclude that the reason for setting exams does not apply in online courses, as we cannot be at all sure that the work is the student's own, or that all students in a course undertake their exams in comparable conditions.

Is there any point in conducting assessment in remote online courses?

By the same token, we must wonder if there is any point in setting any assessment tasks in online courses. If even the exams are not guaranteed to be the students' own work, what hope is there for the assignments, which the students have weeks to complete?

Is there any point in conducting remote online courses?

With such a cynical outlook, we must surely conclude that there is no point in conducting remote online courses; that in doing so we are assuredly at times conferring qualifications on students who have not earned them.

But there is, of course, a pressing reason for conducting online remote courses: our university believes that they bring in lots of money. While the academics who work in an institution might well be idealistic,

the institution itself must of necessity be commercial. Ruhe (2002) points out the spectacular growth in online enrolments at institutions in the USA. This trend is not restricted to that country, and many institutions are reliant on such enrolments to help them remain properly funded. Indeed, the belief appears to be that if an institution allows itself to fall behind, it will also fall by the wayside.

It is not universally accepted that universities make money through online teaching. In his essay on 'digital diploma mills' (Noble, 1997) David Noble wrote: 'Experience to date demonstrates clearly that computer-based teaching, with its limitless demands upon instructor time and vastly expanded overhead requirements – equipment, upgrades, maintenance, and technical and administrative support staff – costs more not less than traditional education, whatever the reductions in direct labor.' However, this essay was written some eight years ago, principally about online education for on-campus students; institutions today, if they are aware of it at all, presumably dismiss its arguments as either wrong or no longer valid.

Is there any point in conducting courses?

The logical extension of the previous question is whether even face-to-face courses meet the academic ideal for which they were created. We are aware of students who have passed face-to-face courses not by academic achievement but by bullying and blustering and/or by cheating. We like to believe that such students are in a small minority, and that the work we do really is worthwhile. But if we believe that for face-to-face students, should we then persuade ourselves that it is true for remote students as well? Should we simply stop worrying about the cheats and assess as if everyone were honest? It would certainly be a great deal easier, except perhaps on the conscience.

Are we interested only in detecting cheating?

We have been asked why we do not direct our effort towards student learning, concentrating on preventing cheating in the first place. We reply that we do. The work reported in this paper is one aspect of our approach, and should not be taken as suggesting that catching and punishing cheats are our only goals.

We have been asked why we do not try to alter the way in which teaching and learning happen so that cheating is assumed not to happen, so that cooperative and ethical behaviour is a natural outcome. To this we pose a parallel question: why do the police and government not alter the way in which society works so that crime is assumed not to happen, so that cooperative and ethical behaviour is a natural outcome? The answer to the parallel question is probably obvious. Even if it were possible to reshape society in such a way, it would be a massive undertaking that would last hundreds of generations. But many great thinkers have concluded that it is simply not possible. Of course assuming that there will be crime is by no means the same as assuming that everyone is a criminal. The measures put in place to detect crime and other misdemeanours (speed cameras, red-light cameras, video surveillance, etc) are based on the assumption that *some* people will err, and the desire to detect it when this happens. If anybody finds a way to recreate society so that nobody will ever err again, the measures can be dispensed with.

It would be churlish to condemn the noble ideals implied in these questions, but it would be naïve to assume that any approach based on these ideals would completely do away with cheating. Our system has clearly indicated that cheating does take place in our courses. It is our contention that so long as cheating takes place, it is better to detect it where possible, and to deal with it appropriately, than to turn a blind eye to it.

Conclusion

Alert to the possibility of a particular form of cheating in remote online exams, we devised a system to detect that sort of cheating, and found that it does indeed happen. We are aware of other possible ways of cheating, ways that we have not yet worked out how to detect, so we assume that these other ways of cheating are also being employed.

We do take steps to try to reduce the motivation to cheat, but we are not confident that these steps can ever completely eliminate cheating. We are left wondering whether to enhance our system so that it covers more and more forms of cheating, if indeed that is possible, or to give up, go with the flow, and behave as if there were no such thing as cheating.

References

- Clarke, M., Butler, C., Schmidt-Hansen, P., Somerville, M. (2004). Quality assurance for distance learning: a case study at Brunel University. *British Journal of Educational Technology*, 35(1), 5-11.
- Fröhlich, R., (2000). Keeping the wolves from the doors... wolves in sheep's clothing, that is. *Fourth International Computer Assisted Assessment Conference*, Loughborough, UK.
- Lim, C.P., Hung, D., Wong, P., Chun, H. (2004). The pedagogical design of ICT integration in online learning: a case study. *International Journal of Instructional Media*, 31(1), 37-47.
- Mason, R. (1989). A case study of the use of computer conferencing at the Open University. *Unpublished doctoral dissertation*, Open University, England.
- Noble, D. (1997). Digital Diploma Mills: The Automation of Higher Education. *Essay distributed by Ontario Confederation of University Faculty Associations*.
- Ruhe, V. (2002). Issues in the Validation of Assessment in Technology-Based Distance and Distributed Learning: What Can We Learn From Messick's Framework? *International Journal of Testing*, 2(2), 143-159.
- Sheard, J., Carbone, A., Dick, M. (2002). Determination of Factors which Impact on IT Students' Propensity to Cheat. *Proc. Fifth Australasian Computing Education Conference*, Adelaide, Australia, 119-126, ACM Press.
- Simon (2005). Electronic Watermarks to Help Authenticate Soft-copy Exams. *Proc. Seventh Australasian Computing Education Conference*, Newcastle, Australia, 7-13, ACM Press.
- Thomas, P., Carswell, L., Price, P., Petre, M. (1998). A holistic approach to supporting distance learning using the Internet: transformation, not translation. *British Journal of Educational Technology*, 29(2), 1-13.
- Thomas, P., Price, B., Petre, M., Carswell, L., Richards, M. (2001). Experiments with electronic examinations over the Internet. *Fifth International Computer Assisted Assessment (CAA) Conference*, Loughborough, UK.
- Thomas, P., Price, B., Paine, C., Richards, M. (2002). Remote Electronic Examinations: student experiences. *British Journal of Educational Technology*, 33(5), 537-549.
- Thorpe, M. (1998). Assessment and 'third generation' distance education. *Distance Education*, 19(2), 265-286.
- Zobel, J. (2004). "Uni Cheats Racket": A Case Study in Plagiarism Investigation. *Proc. Sixth Australasian Computing Education Conference*, Dunedin, New Zealand, 357-365, ACM Press.

Copyright © 2005 Simon: the author assigns to HERDSA and educational non-profit institutions a non-exclusive licence to use this document for personal use and in courses of instruction provided that the article is used in full and that this copyright statement is reproduced. The author also grants a non-exclusive licence to HERDSA to publish this document in full on the World-Wide Web (prime sites and mirrors), on CD-ROM, and in printed form within the HERDSA 2005 conference proceedings. Any other usage is prohibited without the express permission of the author.