Work in Progress - Reuse of Homework and Test Questions: When, Why, and How to Maintain Security?

Edward F. Gehringer
Dept. of Electrical & Computer Engineering, Dept. of Computer Science, North Carolina State University
Raleigh, NC 27606 efg@ncsu.edu

Abstract - It is always difficult to obtain good homework problems and test questions. Instructors can save time—and polish their work—by using the same questions they used before. And they would do so semester after semester, except for one obvious risk: that students would simply copy or memorize the answers rather than learning the material. This paper presents the results of a survey of hundreds of postsecondary educators. How frequently do they reuse questions, and how do they prevent students from getting advance access to the answers? How much trouble have they had with “files” kept by fraternities, sororities, and ethnic groups? Do they consider it cheating to copy or memorize answers? Has the increasing use of electronic resources made it easier or harder to maintain security? Do they typically alter their policies when they begin to put material on line? The answers to these questions can guide all of us to more realistic and secure reuse policies.

Index Terms - Homework, tests, examinations, reuse, cheating, academic integrity.

1. THE SURVEY

Devising new homework assignments or test questions is hard work, and it is not likely to be rewarded in any way. Thus, one would expect faculty to do as little of it as possible, reusing questions rather than creating new ones. But how often do they reuse their own material? There seems, however, to be little if any published work on this topic. The author created a Web-based survey on this topic and announced it on several listservs, including the ASEE Engineering Technology listserv, the ACM Special Interest Group on Computer Science Education member listserv, and the NACADA Commission on Technology in Advising TECADV-L list. We later solicited instructors individually, as described below. The survey was conducted from January to April 2003. Three hundred forty responses were received. Canada furnished 9, Australia 8, Taiwan 5, and Turkey and the UK sent one each. Of the 340 responses, 234 came from engineering educators. Of these, 91 came from computer and/or information scientists. Because of the fact that some CS departments are in engineering and some are not, totals for engineering are reported two ways, with and without computer science. The responses from the “other” fields were about equally divided between science, agriculture, and the humanities, with a scattering from other fields. Anonymous responses were allowed, but instructors were allowed to give their name and e-mail address for followup questions. Instructors were allowed to respond multiple times, once for each different level of course they taught. Only a few did, though the exact number is not discernible due to anonymous responses.

Although we posted our survey on several listservs, relatively few of their ≈ 3500 members responded to the survey. One might therefore question whether the responses were typical of the experiences of faculty across the country. We therefore devised a methodology to survey faculty “at random.”

We used two lists: A list of American universities, http://www.clas.ufl.edu/CLAS/american-universities.html, and a list of engineering graduate schools, http://www.allaboutgradschool.com/usgradschools/engineering/engineering.htm. We used the first list, sorted it randomly (based on Excel's random-number function), and took the first few dozen schools on the list. Then we tried to select 3 engineering departments and 3 other departments from the school. We looked up e-mail addresses of all faculty in these departments, and chose every third one to receive the survey. However, this approach yielded very few engineering faculty, so we then went back and used the list of engineering graduate schools to try to balance it, following a similar procedure. According to this methodology, we sent the survey out to 656 faculty. Only 79 of these answered the survey, a response rate of 12%. Although this was disappointingly low, it was still quite a bit higher than the rate at which listserv members answered, and it was not biased toward faculty who are on teaching or advising listservs (whom one might think would be more attentive to teaching issues).

2. THE EXTENT OF REUSE

Data for the frequency of reuse is given in Table I. Notice that homework assignments are, on average, reused more frequently than exam questions. They are used every year to year-and-a-half, whereas exam questions are reused every two years. The longest reported period between reuses is five years for homework and programs, and ten years for exams. There is no major difference between academic fields, though
engineering instructors seem to reuse homework slightly less frequently.

<table>
<thead>
<tr>
<th></th>
<th>Homework</th>
<th>Programs</th>
<th>Exams</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td># who</td>
<td>Time</td>
<td>Time</td>
</tr>
<tr>
<td></td>
<td>reuse</td>
<td>uses</td>
<td>uses</td>
</tr>
<tr>
<td>Engineering w/o Comp. Sci.</td>
<td>142</td>
<td>112</td>
<td>1.33</td>
</tr>
<tr>
<td>Engineering, incl. Comp. Sci.</td>
<td>232</td>
<td>173</td>
<td>1.35</td>
</tr>
<tr>
<td>Computer Science</td>
<td>90</td>
<td>61</td>
<td>1.35</td>
</tr>
<tr>
<td>Neither Eng’g nor Comp. Sci.</td>
<td>108</td>
<td>62</td>
<td>1.24</td>
</tr>
<tr>
<td>All</td>
<td>340</td>
<td>235</td>
<td>1.32</td>
</tr>
</tbody>
</table>

Data for instructors who reported they reused assignments. Time given in years. Totals for engineering given with and without computer science included. The data do not report how many instructors did not reuse assignments; of those who left the question blank, some but not all reported that they did not reuse.

3. FACTORS ALLOWING REUSE

3.1. Student isolation

Instructors who frequently reuse their assignments have several rationales for why this does not give some students an unfair advantage. Some simply say that their students do not communicate with each other. As one instructor put it,

“In a community (2-year) college there is no real networking of students from one semester to the next.”

Several other community-college instructors concurred, as did one instructor from a “non-residential branch campus.” Contrast this with the statement of an Engineering Technology instructor at a private urban university in the Midwest:

“Our students have shown a significant tendency to keep exam files, homework files, project files. There are three fraternities/sororities on campus that keep files. For this reason, I would never consider using test questions or exam questions a second time. I often turn them into homework problems by changing the wording a bit. They don't seem to have caught on to that yet.”

A Mechanical Engineering Technology instructor at a large public urban university takes this into account when weighting answers:

“All of my questions require the student to show the process by which the answer was derived. Correct answers by themselves are worth 10%, the process the rest.”

However, to state the obvious, this does not guard against students communicating the process behind an answer as well, especially since the student must write it on the test paper.

3.2. So what?

The second rationale might be called the “shrug approach”: If students copy, so what? There are several ways of rationalizing this.

- Homework doesn’t count for much of the grade anyway. In the words of a Math and Computer Science instructor from a small private college in the East:

  “The homework problems in the classes I teach are typically not turned in, we discuss them in class. So reusing these aren't [sic] a problem.”

Other instructors said that they counted the homework for only a small fraction of the grade, e.g., 10%, so they did not worry about copying.

- Students don’t perform any differently on reused questions than new questions, so one may as well reuse questions. Two instructors mentioned this. As a Computer Science instructor from the Midwest put it,

  “Teaching for 25 years, I have never seen any significant statistical difference in student’s performance when given old questions versus new ones no matter what time frame intervened.”

- Copying doesn’t need to be policed, because students who copy are only hurting themselves. As a professor of Textile Technology Management from the South says,

  “Believe strongly that good students will use the exam questions and answers to learn material. So am more interested in enabling the good student learn pre and post exam than protecting against the potential reuse of the answers by students with no standards.”

- Even if they just look at others’ homework, students will still learn something. An Information Technology instructor at a large church-related university in the West says—

  “I insist that homework show HOW they arrived at their answers. If they learn how to get the right answers by looking at the homework from past years, it doesn't matter, so long as they write it out themselves. A great deal is learned that way, too.”

- Using all new problems still won’t guarantee academic integrity. A Civil Engineering Technology instructor finds that—

  “I don't worry about repeating homework assignments because students can copy others' work regardless of how diligent you are about changing the problems. Many students don't do all of the assigned homework anyway and I have recently resorted to collecting homework unannounced about three times in a
week semester and counting it 10% of their grade. Those who elect not to do homework on a timely basis earn a grade level less in the course as a result of losing 10% of the points toward the final grade.

Comment: The author finds all of these justifications rather specious. Nothing will totally prevent cheating, just like nothing will totally prevent murder, but that doesn't absolve us from taking a variety of measures to reduce the incidence.

Students who learn by cheating may be learning something different from what we want them to learn, e.g., how to retrieve information instead of how to derive it. They do harm themselves, but they also harm their future employers and the reputation of the school from which they graduated.

3.3. Is copying really cheating?

However, in some environments, one might question whether copying is really cheating. If all students have full access to old assignments, no one gets an unfair advantage from copying the answers to reused questions. A computer-science instructor from a regional public university in the South said,

“I try not to maintain security. I assume that some students will have access to previous materials and I try to give the others equal access.”

This might be termed a generalization of the philosophy behind open-book exams. In the words of an engineering instructor at a state university in the West:

Why should I care if they've seen the solutions before?

My exams are open-book, open-notes, whatever else they want to use -- as long as they don’t communicate with another person during the exam. What's the big deal?

When they get out on the job, their manager isn't going to want to use -- as long as they don't communicate with another person during the exam. What's the big deal?

One might counter that this just teaches students how to copy, not how to do the problems. However, this is not necessarily the case in the case of copious test questions presented without solutions. In this case, it might well be easier to learn the material. One engineering instructor related—

“I went to a school where all exams ever given in a course were available in the library. Any student who did all the problems from all the previous exams deserved an A.”

One good way to encourage studying is to promise to reuse a question or two. A computing instructor from a public university in Eastern Canada says,

“I normally do not recycle test or exam questions, but occasionally I have given students lists of practice problems from previous exams and told them ‘one of these will appear on your midterm or final exam.’”

3.4. No alternative?

Finally, some instructors argue that because of the nature of their material, they have no alternative but to reuse questions:

“In History courses, I don't think that reusing exam questions compromises the course. Unless one changes the focus or themes of the course, which provide the organizing principles behind the lectures and course readings, the students will inevitably be asked the same or similar questions on the exams.”

“This is a beginning welding course. The material is very basic; safety, for example. How many ways can you ask when are you supposed to wear safety glasses? (at all times!) Unlike math-based courses, you can't just change the numbers to get different answers. And there's not a lot of material in a 2-credit course that is half lab work.”

4. The Extent of Cheating

The survey asked instructors how much trouble instructors have had with students turning in various kinds of work that they have copied (or in the case of exams, memorized) from another source. Results are shown in Table II. It can be seen that cheating on homework is perceived to be a much greater problem than cheating in the closed environment of an exam.

<table>
<thead>
<tr>
<th>TABLE II PERCEIVED EXTENT OF CHEATING</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>Perceived extent</td>
</tr>
<tr>
<td>------------------</td>
</tr>
<tr>
<td>Extreme (20% of students or more)</td>
</tr>
<tr>
<td>Very great (10%-20% of students)</td>
</tr>
<tr>
<td>Great (5%-10% of students)</td>
</tr>
<tr>
<td>Occasional (2%-5% of students)</td>
</tr>
<tr>
<td>Rare (maybe once a year)</td>
</tr>
<tr>
<td>Very rare (once in several years)</td>
</tr>
<tr>
<td>I have never observed this.</td>
</tr>
</tbody>
</table>

Incidence of cheating perceived by instructors. Not all instructors answered all questions; those who did not answer were excluded from the calculations.

5. Defenses Against Cheating

5.1. Controlling Distribution.

The survey suggested several techniques for defending against unauthorized student use of reused questions. Many of the techniques involved controlling distribution of the questions and/or answers. The most popular technique, mentioned by 46% of the respondents, was not posting answers on the Web. It is fairly obvious that one should not post the answers to textbook questions on the Web, lest their usefulness to others be destroyed. However, posting answers to one’s own questions raises the risk that students will store them electronically, and later disseminate them. Of course this can also be done with paper solutions, but it is easier if they are in electronic form.

Posting material in a restricted area of the Web, visible only to the students in the class, is a little more secure than posting them in an unrestricted area. About 18% of our respondents reported that they posted answers in a restricted area of the Web. Almost the same number said they posted assignments in a restricted area of the Web. The responses suggest that many of these use a course-management system.
(such as Blackboard or WebCT) that makes it easy to restrict access, and they do not go to the trouble to place assignments in a different area than answers.

The other techniques for controlling distribution were non-electronic. The most common, employed by 37% of the respondents, was to allow students to see, but not keep, their graded exam paper. This works especially well for final exams, where the graded papers will not be needed to study for anything later in the course. According to several reports, very few students come back to look at their final exams.

Four percent of the instructors said they don’t even allow their students to see their graded exams. At first glance, it seems that this might put the students at risk of being hurt by grading errors. However, the only two instructors who elaborated on this strategy said they used it for multiple-choice exams only, which can be graded by computer. Five percent of the instructors said they place exam answers on a glassed-in bulletin board, so students must copy them manually. These five instructors come from widely different geographic areas and different academic fields.

Respondents mentioned several other approaches that were not listed on the survey, including
- not allowing anything, even scratch paper, to be removed from the final-exam room,
- allowing students to view their midterm tests only in the instructor’s office and get help at that time,
- just reviewing exams in class, rather than passing out answers, and
- removing answers from the Web soon after assignments are returned to students.

5.2 Random variations in problems

The second major defense against cheating was random variations in problems. This is easy using a Web-based assessment and testing system such as LON-CAPA [1], Mallard [2], or WebAssign [3]. Though this approach was not listed on the survey, 4% of the instructors reported using it. These systems allow an instructor to vary the numeric inputs to a problem randomly for each student, and then calculate the answer and compare it against the answer submitted by the student. Even if the problem is not numeric, it may be possible to change the inputs in a straightforward way. A Computer Science instructor at a selective teaching-oriented university in the South wrote,

“Use parameterized homework and test questions, [e.g.,]
prove that the following grammar is ambiguous (where
the grammar varies each semester).”

Long before the Web, a geneticist employed a Snobol-4 “Simple Mendelian Generator” program [4] to randomly select traits from a preset list and combine individuals with these traits in various ways, and give the observed outcomes in their offspring. The student was asked to determine whether certain genes were dominant, recessive, etc. There are undoubtedly many other examples of special-purpose programs for generating randomly varying problems in other fields.

5.3. Software to detect plagiarism

Recently there has been a lot of interest in computer programs to detect plagiarism, both in prose (e.g., Turnitin.com [5]) and programming (Moss [6]). These systems were used by 8% of our respondents. They theoretically make it possible to give the same writing or programming assignment semester after semester without risk of undetected plagiarism, though they require a certain amount of effort to configure and interpret the results.

5.4. Some caveats

As several respondents noted, some techniques for controlling dissemination of answers may have undesirable effects on student learning. A computer scientist at a historically black state university put it this way,

“While the above techniques improve the security of the material, they reduce the teaching quality. It is useful for a student to see the correct answers. If they see the answers in any form, they can copy them.”

Another computer scientist noted that the honest learners should not be penalized for the actions of the dishonest.

“[T]he real learners are not cheating anyway,” he said.

In an ideal world, students would get immediate feedback. As a Civil Engineering Technology instructor at a large urban university said,

“Students must have immediate access to answers to the exam questions - the sooner after the exam, the better! The exam should be part of the learning process. Students should walk away from the exam with a written solution.”

Web-based testing systems can give both immediate feedback and security for numerical problems, but not necessarily for other kinds. Nonetheless, they are probably still better than other means of question delivery, according to this Business Management instructor:

“It does not seem like good pedagogy to prevent students from seeing and retaining answers to questions. The answers are an important type of feedback for them. The best solution is to vary questions frequently and/or to provide the answers in a form that is unlikely to be copied. I worry somewhat about answers to discussion questions on WebAssign. However, I assume that students are unlikely both to copy the answers and then to pass them on to future students in the class. Doing so would require too much work.”

6. DOES FREQUENT REUSE LEAD TO CHEATING?

Does more frequent reuse of questions lead more students to cheat? We ran a variety of statistical tests to try to establish significance. Significance could not be established for reuse of homework. For reuse of test questions, we employed two categorical methods. One of these was nonzero correlation (also called the linear-trend statistic). This test compares two variables, which are both intervals, in this case, the length of time between reuse and the level of perceived cheating. The
second is “row mean scores differ.” It compares a nominal variable (without a natural ordering) to an interval variable (which has a numerical distance between any two values).

When these two tests were run, the $p$ values (0.2077 and 0.3468 respectively) were outside the range of significance. However, three instructors gave anomalous answers, saying that they never reused questions, but nonetheless had problems with cheating from students who copied or memorized answers to previous tests. When these three responses were removed from the data, both tests showed a positive correlation significant at the 90% confidence level between more frequent reuse and the amount of cheating observed (Table III).

<table>
<thead>
<tr>
<th>Alternative hypothesis</th>
<th>W/o dropping anomalous data</th>
<th>Dropping anomalous data</th>
</tr>
</thead>
<tbody>
<tr>
<td>DF</td>
<td>$p$</td>
<td>$p$</td>
</tr>
<tr>
<td>Nonzero correlation</td>
<td>1</td>
<td>0.2077</td>
</tr>
<tr>
<td>Row mean scores differ</td>
<td>15</td>
<td>0.3468</td>
</tr>
</tbody>
</table>

What about the anomalous answers we removed from the data? We asked the instructors to explain them, and received responses from two of the three. One instructor said that she was thinking of a case where they had discovered answers written during the exam and left in a restroom for other students to read; about six students admitted being involved in this scheme. This was not a true case of students copying answers from an earlier exam, and thus was correctly removed from the data. The other instructor said that his students “give the right answer to the old questions.” Again, this was not a case where students cheated; they tried to cheat, but failed.

One possibility is that instructors who notice copying of answers from earlier exams adjust their reuse policies to minimize it, and therefore that our data understates the extent of the correlation between frequent reuse and cheating. The current data do not provide a way to determine whether this is the case, but a followup study is being considered.

7. INCREASING THE SUPPLY OF QUESTIONS

If questions are to be reused less frequently or not at all, faculty must have access to a supply of new questions. Survey respondents thought this quite important (Table IV). The survey suggested several ways of obtaining more questions. The most obvious—and also most common method—is to write them oneself; 86% of respondents reported doing that.

One computer-science instructor noted,

“I dislike preparing questions and assignments, but I dislike using those prepared by others even more (in my experience it's just as much work to adapt someone else's ideas as it is to make up my own).”

Next most common was taking questions from textbooks other than the textbook for the class; 61% said they did that. According to an Information Technology lecturer from Australia,

“I’ve found that with the ever-changing technology, there is a continuously changing supply of texts which gives me sufficient ‘new’ materials.

After that came taking questions from the Web, reported by 24%, and borrowing them from colleagues at one’s own institution with their permission (23%). Relatively uncommon approaches were to borrow questions from instructors at other institutions (13%), or have one’s teaching assistants write them (9%). One Engineering Technology instructor noted that with LON-CAPA, it is easy to use questions written by others who also use the system (the author has heard that the same is true of Mallard). One instructor at a California university said that now that their teaching assistants were unionized, he rarely has them write questions.

Other sources mentioned by respondents were—

- test banks that come with textbook adoptions (Information Systems Technology instructor),
- industrial contacts,
- students, and
- teaching-oriented conferences.

Among those who mentioned industrial contacts, an Engineering Technology instructor said that she had the people from industry write the questions:

“I have often asked industry professionals to help me write exam and programming questions. They really seem to enjoy it, I learn a lot, and the students get real-life based questions. Later I reuse these questions in-class for in-class work and discussion. I reward helpful professionals by taking them to lunch.”

Instructors who have industrial experience can draw on it, like this Textile Technology Management instructor:

“Teach mathematical and modeling methods from theoretical point of view. Then have homework questions set in practical industry settings. Then ALL exam questions are couched in a practical, industrial setting, including extraneous data and information, that challenges student to interpret situation, select best problem solving method and then execute method effectively and accurately. Approximating my experience with general examinations for ChE PhD or what I faced in industry for 30 years.”

Students can be assigned to come up with problems as part of their coursework, and some of these can later be used. One Computer Science instructor said,
“Have the students submit 3 questions they feel would be good ones for upcoming exams. Any that are used get extra credit for the student(s) that submitted them.”

Another instructor uses student-written questions on the final exam; any student whose question is chosen is thus “advantaged.” The author has his students submit questions for peer review [7], and uses the best of them for later homework and exams; in some courses, up to 25% of the questions are obtained via this route.

The last source, teaching-oriented conferences, was mentioned by only two respondents. Both of them referred to the Nifty Assignments panel at the SIGCSE Technical Symposium on Computer Science Education. One of the respondents said he “would like to see an online repository of questions categorized and accessible only by instructors, with content coming from whoever is interested in supplying material.” The author has developed a prototype of such a database, which is described in another paper [8].

8. SUMMARY

Postsecondary instructors were surveyed about their practices in reusing homework assignments and exam questions. Most, but by no means all, instructors do reuse their own material. Typically they use the same questions every year or two. Although they employ a variety of mechanisms to maintain the security of exam questions, some of their practices with respect to homework seem a bit naïve. Increasing frequency of reuse of exam questions (but not homework assignments) seems to be positively correlated with the amount of cheating that instructors observe; once a small amount of anomalous data is removed, this finding is significant at the 90% confidence level.

ACKNOWLEDGMENT

The author wishes to than NCSU Statistics graduate students Clay Barker and Jianyuan Zhu for performing the data analysis reported in Section 6. Teaching assistants Bensong Chen, James Gritter, and Sree Vidhya Lakshmi Krishnakumar assisted in gathering names of instructors who could be surveyed.

REFERENCES