

Multiple choice questions to combat plagiarism and encourage conceptual learning

Gareth Denyer and **Dale Hancock**, School of Molecular and Microbial Biosciences, The University of Sydney, Australia

gareth@mmb.usyd.edu.au d.hancock@mmb.usyd.edu.au

Plagiarism in assignments and examinations

Academic dishonesty (i.e., plagiarism or cheating in plain English) discredits and compromises the validity of University qualifications. All the effort an academic puts into the construction, administration and delivery of their courses is undermined if students are cheating.

There are a number of studies describing methods for detecting plagiarism in written assignments. For example, there is software which can detect high levels of similarity between students' work (Moeck 2002; Hamilton 2003), and there has been a special interest in identifying Internet-related cheating (Young 2001), commercial ghost-writing (Hammer 1976) and in alerting academics to other forms of electronically based plagiarism (McMurtry 2001; Heberling 2002).

At first glance, these issues seem less relevant to our large Biochemistry and Molecular Biology classes (which have enrolments of between 200-700 students) because our assessment is based on laboratory work (including individualised practical skills tests) and end-of-semester examinations. However, not only is it clear that a great deal of cheating can occur in the examination room (especially when large numbers of multiple choice questions are employed), our experiences have also made us realise that the detection of cheating causes tremendous stress of itself.

Thus the central theme of this paper is that it is a waste of time to employ sophisticated plagiarism detection solutions for in-semester written work if those assignments contribute relatively little to the final mark (as would be the case if the bulk of the final assessment is an easily corrupted multiple choice examination). Additionally, we believe that it is better to make multiple choice examinations that cannot be plagiarised than to burden oneself with the problem of dealing with suspected plagiarism. To this end we have developed examination solutions that prevent cheating and which, as a by-product, allows for flexibility in question design that facilitates the grading of individual options within multiple choice questions.

The problems with plagiarism detection solutions

The battle between the plagiarist and the educator is a 'cat and mouse' game (Young 2001) akin to that between the email spammer and the IT-systems administrator. As the latter constructs more intelligent solutions to detect unsolicited mail, the former adopts new strategies and tricks. All this leads to a lot of time, effort and expense being spent on reactive remedies which can never be 100% reliable.

A more serious problem is the question of what you actually do when you detect plagiarism. The investigation process is upsetting for all parties. Students (and, indeed, their parents) have been offended by our accusations or investigations. This is especially true when one party is innocent (i.e., when one student has cheated off another without their knowledge or consent) because, despite studies like Ross (1977) which show that it is usually the student with the lowest grades that has cheated, proper process dictates that both students are subjected to the same rigor of interrogation and suspicion.



An investigation into alleged plagiarism is not pleasant for academics either. Fundamentally we are teachers, not policemen. Instigating an official investigation into a student's work (with potential disciplinary action) creates a wide and often unbridgeable rift between us and our class. The investigation process is administratively time-consuming and emotionally draining. The outcome is rarely seen as satisfactory by the academic because University administrators are justifiably cautious in their investigations and ultimate decisions. The possibilities of legal escalation (Mawdsley and Permuth 1986) weigh heavily on the decision maker's mind and, despite some researchers recommending that future offenders could be deterred by exposing the essence of on-going enquiries (Todd-Mancillas 1987) administrators have to weigh up the potentially good and bad publicity that could result from releasing their efforts to address plagiarism. Every time a student escapes penalty, however, academics become less likely to report cases of suspected plagiarism and more cynical and disenchanted with the Administration. Indeed, students are apparently well aware of academics' tendency to not make too much extra work and stress for themselves (Love and Simmons 1997) in this regard. It is not implausible to imagine that the prospect of having to detect and act upon plagiarism is a factor in deterring academics from setting in-semester assignments, and in our continued preference for relying on end-of-semester examinations. However, it should be appreciated that even those examinations are not fraud-proof.

Cheating in multiple choice examinations

It might come as a shock to many academics to know that multiple choice question (MCQ) based examinations are highly susceptible to corruption. Yet there is a rich body of literature that reveals many forms of cheating in these tests - from casual or opportunistic through to organised and pre-meditated (Bellezza and Bellezza 1989; Frary 1993; Sotaridona and Meijer 2001; van der Linden and Sotaridona 2002). Given the usually high contribution of MCQ examinations to a student's final mark (especially in science), it seems strange that so much emphasis has been put on the detection and/or elimination of cheating in hand-in assignments.

But how can this MCQ examination cheating occur? After all, academics will have done their best to keep their manuscripts confidential, and we know that staff in the Examinations Office are rigorous with their processing and distribution of the exam. Additionally, as any academic who has tried to enter an examination room would testify, invigilators are suspicious and protective to the point of paranoia.

But, in fact, it is rather easy to cheat in an MCQ examination. In many of the large examination halls students sit within 1-2 m of each other and, whilst this distance might prevent the clear viewing of handwriting, it is no impediment to the copying of MCQ answers. Remember that in an MCQ test, the students' answers are recorded on a computer answer sheet (that contains 50 to 100 answers per page) by filling in a grid of blank circles with a dark pencil. A student with wandering eyes does not even have to look at the question booklet to get the answer; they merely copy the "pattern" of answers off their neighbours' computer sheet. The closeness of the students' tables and the ease of identifying patterns on the rigidly structured MCQ answer sheets make casual and opportunistic cheating irresistible. Indeed, we often see patterns that indicate that a student has copied each page of the answer sheet, or separate blocks of questions, from two or more individuals. However, perhaps a more serious problem is that predictable seating arrangements make premeditated, collaborative cheating straightforward (Houston 1986). Even cheating between students that are sitting several places apart is possible if pre-arranged hand (or other) signals are rehearsed (although the evidence that this occurs is purely anecdotal and is not likely to account for most of the misconduct in an examination).

Statistically it is quite straightforward to detect that cheating has occurred in MCQ examinations (see references above). However, it is much harder (if not impossible) to prove which student is the



cheater and/or if the cheating was collaborative (Roberts 1987). More importantly, and as discussed in the previous section, it is better if the time, expense and energy was not wasted on the detection and reporting of plagiarism, but that instead, the cheating was just rendered impossible. Some attempts have been made to do this in the past, notably the multiple numbers solution of (Tauber 1984) and the multiple ordering system of (Klein and Bolus 1983), but such solutions require specialised software. So, in the late 1990s, we developed an anti-cheat multiple choice system that uses basic word processing and spreadsheet software (e.g., *Word*, *Excel*, etc) for creation and marking of manuscripts.

The anti-cheat MCQ solution

Design

Details of our solution, including the instructions on construction and implementation, are described in (Denyer and Hancock 2002). In brief, a standard MCQ examination is prepared in *Microsoft Word* and, from this, four separate versions of the paper are created by simply rotating the options within each question. This is made easy with the Word 'drag and drop' and renumbering features. The question order within each version of the paper is kept the same, but the option order differs between the versions. It is possible to lay out the papers in the examination hall such that each student is surrounded by colleagues attempting a different version of the exam. Thus all the answer sheets within viewing distance of any individual student relate to a different version of the paper to the one that they are sitting. If a student copies off any of their neighbours' correct answers then, by definition, they choose the wrong answer on their own sheet. Note that the system makes it easy to determine exactly who has cheated off whom. Although the latter is made all the more easy by our insertion of 'check' questions, our major aim is not to detect cheating but to simply penalise it with natural justice. The completed answer sheets are processed normally, except that there is a separate key sheet for each version.

Implementation

We have been running the scheme for about seven years and have implemented it in over twenty large MCQ examinations. Although we find it necessary to assist the invigilators in the layout of the examination hall, this is actually a good thing as it gives us a deeper appreciation of the examination experience for both invigilator and student. Indeed, we would recommend that academics take the time to attend their examinations. It is good for student morale to see their instructors taking an interest in the examination and allows for rapid correction to any during-examination problems.

Although there have been some concerns that the pattern of options within a question can alter student responses (Bresnock 1989), those studies only showed a minimal effect, even with questions that were carefully constructed to lull students into a particular train of thought. Indeed, during implementation with several thousand students, we have not observed any significant difference in performance between versions. For example, in 2005 MBLG1001 final examination the distribution of marks (% +/- SD) were: Version A 50.4 +/- 13.8 (n = 158), Version B 51.3 +/- 15.2 (n = 159), Version C 48.3 +/- 13.6 (n = 144) and Version D 52.0 +/- 14.6 (n = 133).

Turning the tables on plagiarists

Currently we do not tell students that they are each sitting a 'frame-shifted' version of the paper to their neighbour. We are aware that this could be seen as entrapment by some, but this is not the motive. Rather, we do not want the scheme to be ultimately compromised by the natural inclination of human beings to develop counter strategies (i.e., we do not want to get into the game of 'cat and mouse' like the standard plagiarism detection packages). This also explains why we do not want to reveal the solution by exposing the cheats. Clearly, this policy is something that we expect to be rigorously discussed. However, there will not be many academics who don't enjoy the following scenario (inspired by personal experience and not a verbatim account).



Student friends X and Y arrive in the academics office. X has scored 90%, Y has scored 10%.

Student Y says: "I want to see my exam paper and get a breakdown of my marks... I just can't understand how I did so badly"

Student Y thinks: "There must be a mistake, I copied the whole paper of my friend X and they got 90%."

Academic says: "Sure" (smiles)

Academic thinks: "I know that they cheated, and they know they cheated but they can't admit it. No investigation. No accusations. No stress. Just justice."

Academic shows student their choices in comparison to the mark scheme.

Student Y thinks: "But it's not fair! I cheated! I should get 90%. But what can I say?! I can't let them know about their mistake without admitting that I copied and that my friend was in on it!"

Rewarding deep learning

Although this article has thus far emphasised its anti-plagiarism features, our system also contains spreadsheet solution to award partial marks to individual options in an MCQ. This not only allows examiners to give some credit for near misses, it also allows for post-examination reassessment of mark schemes. The latter might be applicable if the pattern of student responses shows some ambiguity in the question or even if the academic wants to experiment with a more creative way of scaling the marks for the paper. The real benefit, however, is that much more imaginative and extrapolative questions can be set. This is especially true of questions requiring deduction involving several steps of logic or calculations because each of the options can be set to reveal a particular train of thought or a particular mistake by the student. With a non-graded system, the examiner is much more likely to be conservative and set a non-ambiguous question involving recall of surface-learned fact.

Focus-group discussions have revealed that the graded system is very well appreciated by the students. It certainly favours the higher-achieving cohort because they are the most likely to benefit from partial marks for a 'near miss'. Similarly, if they deduce that the answer is one of two plausible options, they feel happier about making their selection and moving on knowing that they will get some credit for their choice. Conversely, in the non-graded alternative, students can get stuck and fixated on particular questions. Implementation of the graded system is, however, more labour intensive, especially if combined with the anti-cheat rotations.

Writing effective, clear multiple choice questions that test conceptual understanding is very difficult (Marrelli 1995) but it is even harder to convince students that they should take a deep approach to their studies when they know that they will be faced with a 100-question MCQ exam. Again, by showing them the graded questions and mark schemes, we are able to give the students some faith that a deeper approach will be rewarded.

Recent developments

Since (Denyer and Hancock 2002) we have developed a database-driven solution in FileMaker Pro to make the generation and marking even faster and more flexible. Indeed, we can now generate over 100 separate versions of the same paper (there being some 120 different patterns of five options) and, could, if we wished also rotate the question order (as recommended by (Houston 1983) to ensure a totally cheat free solution). Despite this, our experience shows that four versions are adequate to eliminate the effects cheating whilst being well tolerated by Examination Office and invigilators.

Conclusions

Multiple choice questions, whether we like it or not, form a very large component of science assessment. As we have seen, cheating is widespread in traditional MCQ based examinations and the first step in combating this is to make more academics aware of the problem. But making MCQ examinations plagiarism free within the constraints of existing infrastructure is problematic. For example, students' desks could be placed further apart but that would increase the need for venues and invigilating staff. Similarly, cheating could be stopped by running several examinations concurrently in the same room (such that students could not see answer sheets from their course colleagues), but this would be administratively impractical.

In contrast, we believe that our anti-cheat MCQ solution offers a simple, easily administered method for ensuring the academic integrity of examinations. Tens of large examinations have been conducted using the scheme over several years and, over that period, several cheats have got their just desserts without the need for lengthy, costly or stressful investigations. Certainly it seems more logical for coordinators of examination based courses to spend their time fighting against examination fraud rather than trying to fight the smaller and less winnable plagiarism battles associated with minor assessments.

As a further extension, we believe that grading the options in multiple choice questions offers the examiner scope to set more imaginative questions and to promote question-solving strategies in their students that foster deep learning and extrapolation from concepts. In short, our system means that the effort that we put into our teaching is supported by a valid and flexible examination process.

References

- Bellezza, F.S. and Bellezza, S.F. (1989) Detection of Cheating on Multiple-Choice Tests by Using Error-Similarity Analysis. *Teaching of Psychology*, **16**(3): 151–155.
- Bresnock, A.E. (1989) Multiple-Choice Testing: Question and Response Position. *Journal of Economic Education*, **20**(3), 239–245.
- Denyer, G. and Hancock, D. (2002) Graded Multiple Choice Questions: Rewarding Understanding and Preventing Plagiarism. *Journal of Chemical Education*, **79**(8), 961–964.
- Frary, R.B. (1993) Statistical Detection of Multiple-Choice Answer Copying: Review and Commentary. *Applied Measurement in Education*, **6**(2), 153–165.
- Hamilton, D. (2003) Plagiarism: Librarians Help Provide New Solutions to an Old Problem. *Searcher*, **11**(4), 26–28.
- Hammer, G. (1976) Academic Trouble across the Border. *SO - English Quarterly*, **9**, 1-2, 82-86, Spr–Sum 76.
- Heberling, M. (2002) Maintaining Academic Integrity in On-Line Education. *SO - Online Journal of Distance Learning Administration*, **5**(1) Spr 2002.
- Houston, J.P. (1983) Alternate Test Forms as a Means of Reducing Multiple-Choice Answer Copying in the Classroom. *Journal of Educational Psychology*, **75**(4), 572–575.
- Houston, J.P. (1986) Classroom Answer Copying: Roles of Acquaintanceship and Free versus Assigned Seating. *Journal of Educational Psychology*, **78**(3), 230–232.
- Klein, S.P. and Bolus, R. (1983) *The Effect of Item Sequence on Bar Examination Scores*, Rand Corp., Santa Monica, CA.[CIQ74890].
- Love, P.G. and Simmons, J.M. (1997) The Meaning and Mediated Nature of Cheating and Plagiarism among Graduate Students in a College of Education. *ASHE Annual Meeting Paper*.
- Marrelli, A.F. (1995) Writing Multiple-Choice Test Items. *Performance and Instruction*, **34**(8), 24–29.
- Mawdsley, R.D. and Permuth, S. (1986) *Plagiarism and Cheating*.
- McMurtry, K. (2001) E-Cheating: Combating a 21st Century Challenge. *T.H.E. Journal*, **29**(4), 36–38,40–41.
- Moeck, P.G. (2002) Academic Dishonesty: Cheating among Community College Students. *Community College Journal of Research and Practice*, **26**(6), 479–491.
- Roberts, D.M. (1987) Limitations of the Score-Difference Method in Detecting Cheating in Recognition Test Situations. *Journal of Educational Measurement*, **24**(1), 77–81.
- Ross, W.G. (1977). *Cheating on Multiple Choice Examinations: Lead us not into Temptation*. ED147725.
- Sotaridona, L.S. and Meijer, R.R. (2001) *Two New Statistics To Detect Answer Copying*. Research Report, Twente University, Enschede (Netherlands). Faculty of Educational Science and Technology. [BBB31588].
- Tauber, R.T. (1984) *Multiple Numbering: An Alternative to Alternative Test Forms as a way to Reduce Cheating on Multiple-Choice Examinations*. ED252572.



- Todd-Mancillas, W.R. (1987) *Academic Dishonesty among Communication Students and Professionals: Some Consequences and What Might Be Done about Them*. ED296406.
- van der Linden, W.J. and Sotaridona, L.S. (2002) *A Statistical Test for Detecting Answer Copying on Multiple-Choice Tests*. Research Report, Twente Univ., Enschede (Netherlands). Faculty of Educational Science and Technology. [BBB31588].
- Young, J.R. (2001) The Cat-and-Mouse Game of Plagiarism Detection. *Chronicle of Higher Education*, **47**(43), A26-A27.

© 2006 Gareth Denyer and Dale Hancock

The authors assign to UniServe Science 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 this copyright statement is reproduced. The authors also grant a non-exclusive licence to UniServe Science to publish this document on the Web (prime sites and mirrors) and in printed form within the UniServe Science 2006 Conference proceedings. Any other usage is prohibited without the express permission of the authors. UniServe Science reserved the right to undertake editorial changes in regard to formatting, length of paper and consistency.