

## [P18] Improving learning and assessment with confidence-based marking

### Tony Gardner-Medwin

Department of Physiology  
University College London  
a.gardner-medwin@ucl.ac.uk

Confidence-based marking (CBM) has been known for many decades to stimulate constructive thinking by students, and to improve both the reliability and validity of exam data. It is surprisingly little used at present, perhaps due to a degree of scepticism, lack of experience, and the fact that major software vendors have not yet adopted CBM in most marketed products. At UCL we have set up web facilities, using FDTL funding for dissemination of CBM (see [www.ucl.ac.uk/lapt](http://www.ucl.ac.uk/lapt)), to make it easy to experience CBM and to use it in your own institution, with your own material and students. The software was designed with basic sciences in mind – Maths, Physics and Chemistry, Biological Sciences and Medicine. Its first application was for teaching Maths and Physics to students of biomedical science.

We have used a simple and flexible, carefully designed CBM scheme for ten years now, mainly with the aim of improving student study habits, rigour of thought, and depth of reflection. In the last four years we have also used it (partly at the students' instigation) in medical 1st and 2nd year summative exams. The software became known as **LAPT** (London Agreed Protocol for Teaching) following collaboration between several London medical schools now mostly subsumed within UCL and Imperial College. The mark scheme is shown in Table 1. It rewards students for correct answers given with high confidence, while penalising confident wrong answers. The increasing ratio of penalty to reward with higher confidence ensures that the scheme is properly motivating:

the student always benefits by reporting his/her true confidence honestly, whether this is high or low.  $C=1$  gives the best score on average if the probability of being correct is  $<67\%$ , while  $C=3$  is best if this probability is  $>80\%$ . We use the scheme both with computer marked exercises (for example, numerical, multiple choice or true/false) and in exams with Optical Mark Reader cards on which the student marks both the preferred choice (MCQ or T/F) and the confidence (1, 2 or 3) for each answer.

The poster sets out the rationale of CBM and the extent of success evidenced by student evaluations and analysis of large bodies of data from formative and summative assessment. Students appreciate CBM a lot, finding it more searching in identifying their areas of weakness or misconception, and more fair as a mode of assessment. They recognise that unconfident right answers and lucky guesses are not the same and deserve less reward than sound justified knowledge, while confident wrong answers warrant both penalty and (in formative work) careful reflection why confidence was misplaced and attention to explanations on offer. Feedback to students about both individual answers and the overall calibration of their confidence for optimal scores are both important features of our software.

An issue often raised by staff is concern that CBM might somehow favour particular personality types, or one or other gender. Our data shows no evidence for gender differences in the calibration of confidence judgements,

**Table 1: The LAPT Confidence-Based Marking scheme ([www.ucl.ac.uk/lapt](http://www.ucl.ac.uk/lapt))**

Confidence level:	C=1 (low)	C=2 (mid)	C=3 (high)	No Reply
Mark if correct:	1	2	3	0
Penalty if wrong (T/F Q)	0	-2	-6	0

either in formative or summative work, despite a significant tendency for both sexes to be less cautious (and more often overconfident) when working for study and revision than in exams. By the time our students take summative exams with CBM they are well practiced, and they generally match their confidence judgements accurately to their probability of being correct, within the optimal bands set out above. Any students who may be, at the outset when they first experience the system, excessively diffident or confident about their knowledge will be helped through practice and feedback to calibrate these judgements correctly. This judgement – the skill of conveying correctly the degree of reliability of one’s knowledge or inferences – is not only relevant to a marking scheme such as we use, but is critical for effective academic communication in any field. The premium that CBM places on the correct identification of either reasons for confidence in a conclusion or reasons for reservation about it, also helps to encourage the habits of rigour, cross-checking and reflection about questions that we all want to encourage in our students. Exam data show that CBM scores are both more reliable (Cronbach alpha) as measures of student performance and better predictors even of the number of correct answers that students will obtain on different questions (Gardner-Medwin and Gahan, 2003).

Teachers interested in trying out CBM with their students can find on the website ([www.ucl.ac.uk/lapt](http://www.ucl.ac.uk/lapt)) sample exercises in a variety of fields. You can use the software to run (from your own website) your own existing or new material, readily adapted from other formats with translation or authoring tools or

word processing routines. Access for students can be arranged within a VLE, with return of grades from LAPT for storage and handling by the VLE software. UCL will currently assist with such development, and is keen to collaborate in evaluations and the development of new material in other institutions and new disciplines.

## REFERENCES

- Gardner-Medwin A. R. and Gahan M.** (2003) Formative and Summative Confidence-Based Assessment. *Proceedings of the 7th International CAA Conference*, Loughborough University, UK, pp. 147-155 ([www.caaconference.com](http://www.caaconference.com))