Abstract for: **e-Assessment: Which Way Now?**
London Knowledge Lab, IoE, 16th May 2013

Certainty-based marking in self-tests and exams

Tony Gardner-Medwin, Dept. Physiology (NPP), UCL
ucgbarg@ucl.ac.uk Homepage: www.ucl.ac.uk/~ucgbarg

It is important to know which things you don't know. The roll call of those who have most neatly expressed this obvious truth stretches from Confucius to Donald Rumsfeld. Assessments, especially self-tests aimed at challenging knowledge and stimulating learning, need to reward the identification and acknowledgement of uncertainty. Certainty-Based Marking (CBM) does this, motivating and rewarding the student for correctly distinguishing between reliable and uncertain answers. With 3 certainty levels (C=1,2,3) as at UCL (www.ucl.ac.uk/lapt) and in Moodle, correct answers gain 1, 2 or 3 marks and incorrect answers 0,-2 or -6 marks. When uncertain (<67% probability correct) it is best to avoid the risk of negative marking by choosing C=1, while the highest certainty level (C=3) is optimal for answers with >80% probability correct. Misconceptions (confident errors) receive a double negative penalty, reflecting the danger if such "knowledge" is applied. In self-tests the fairness, sense and value of this scheme, encouraging careful reflection, is readily accepted by students. In exams, it gives a measure of performance that is not only markedly more reliable in psychometric terms than accuracy measures, but is also a better predictor of the student's simple accuracy on separate questions on similar topics. Looked at also from a purely common-sense point of view, CBM is a fairer measure of performance in that it properly distinguishes, given students with equivalent accuracy, between those with Rumsfeld's "known unknowns" and "unknown unknowns". It is hard to see any justification for not using CBM more widely in e-Assessment.

Certainty-Based Marking in Self-Tests and Exams

Tony Gardner-Medwin, UCL ww.ucl.ac.uk/LAPT

Degree of Certainty :	C=1 (low)	C=2 (mid)	C=3 (high)	No Reply
Mark if correct:	1	2	3	0
Penalty if wrong:	0	- 2	- 6	0

5 minutes to persuade you:

Makes Sense!
Doesn't require special Questions
Always motivates students to give a careful honest judgement

SELF-TESTS

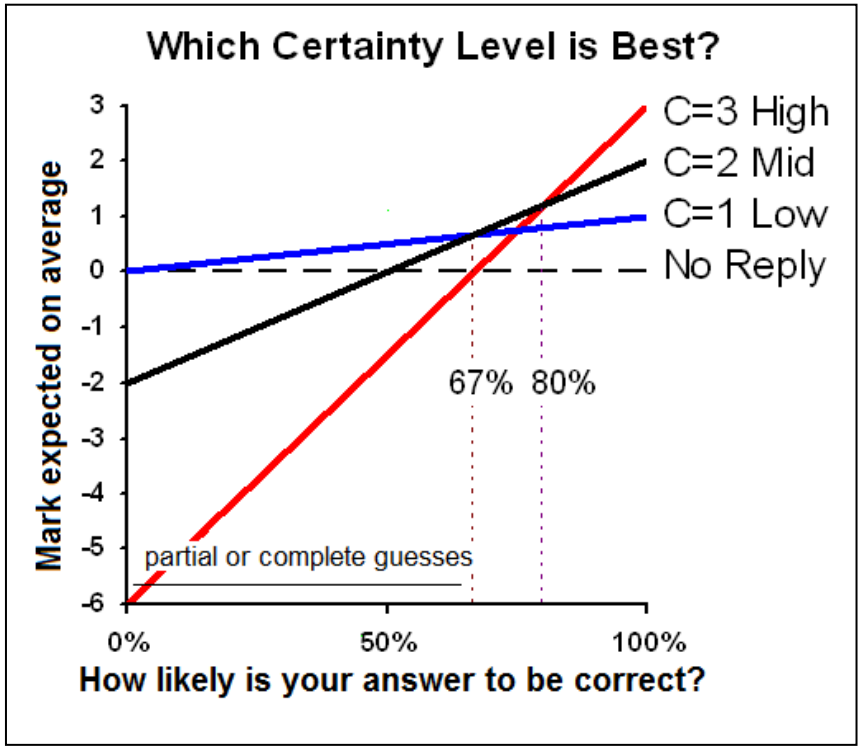
↑ reflection & linking of Info
↑ realism about uncertainty
Highlights misconceptions
Students like it

EXAMS

↑ psychometric reliability
↑ psychometric validity
↓ question numbers
No loss of conventional information

CBM is easy to understand, and motivates honest judgement

Degree of Certainty :	C=1 (low)	C=2 (mid)	C=3 (high)	No Reply
Mark if correct:	1	2	3	0
Penalty if wrong:	0	- 2	- 6	0
Probability Correct:	<67%	67-80%	>80%	–



But I don't like negative marking!
I've never used it in 40 years of teaching!

Fixed -ve marking is seldom rational. It can disadvantage able students, and those who follow advice not to enter guesses.

CBM avoids this -ve marking risk when you are uncertain. It is always best to answer each Q.

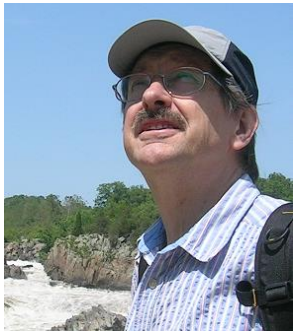
CBM rewards the acknowledging of uncertainty.

"When you know a thing, to hold that you know it,
when you do not know a thing, to allow that you do not know it
– this is knowledge." *Confucius*

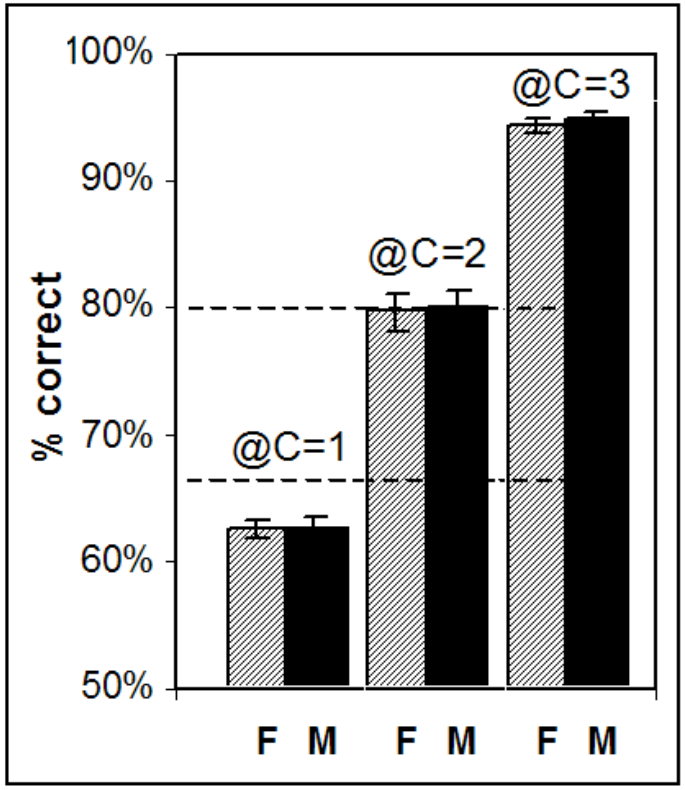
"... there are known knowns;
... there are known unknowns;
... But there are also unknown unknowns *Rumsfeld*

"It's not ignorance does so much damage;
- it's knowin' so derned much that ain't so."
attr.: Billings

"A lucky guess is not knowledge.
A firm misconception is worse than acknowledged ignorance.
So why do we mark students as if these things weren't true?" *TGM*



Students discriminate well



Exams: 331 Sts 500 t/f Qs

Means ± 95% conf. lim.

CBM Self-tests:
what the marks tell you

Very good, but may have repeated self-tests excessively

Good insight into what knowledge is reliable

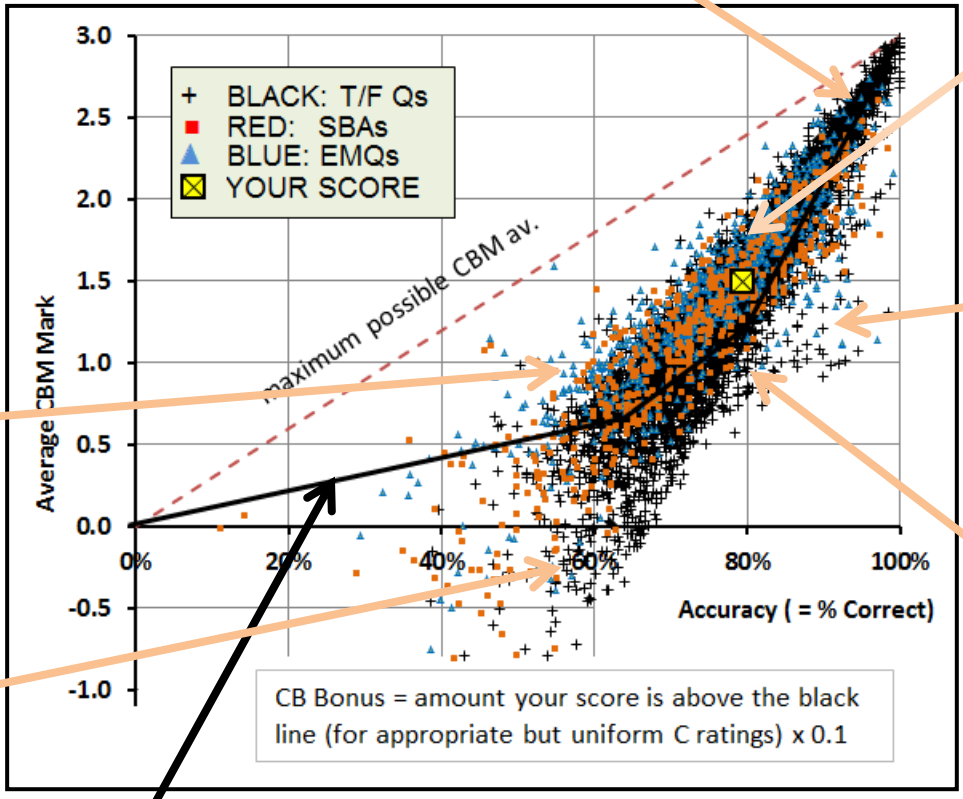
Underestimates knowledge, or not serious about CBM

Knows quite a lot but doesn't know where shaky

Little knowledge but knows what s/he doesn't know

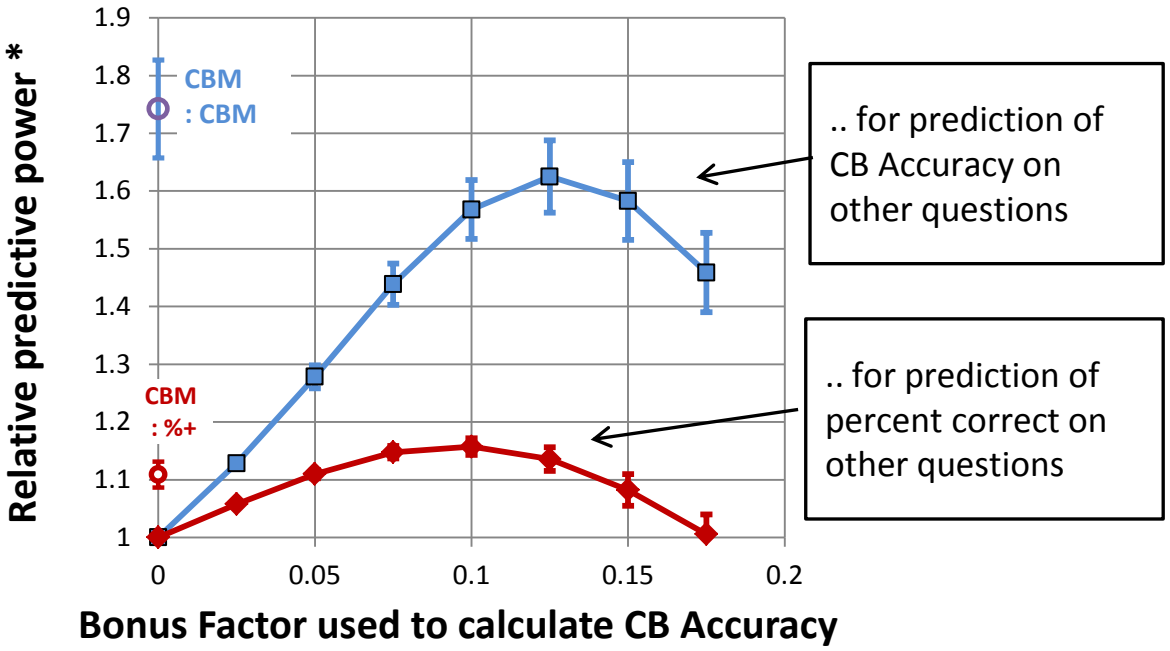
Misconceptions or lack of awareness of ignorance

CBM mark if you use the same C all the time



NB The CBM mark (as a % of maximum) is always bound to be less than the % correct answers

CBM enhances reliability and validity of exam scores



* Factor by which r/(1-r) is increased where r=rank correl. coeff. between scores on odd & even numbered Qs. Mean ± sem for 17 exams, each 250+ t/f Qs, 300+ students.

CBM makes sense!
Doesn't require special Questions
Always motivates students to give a careful honest judgement

SELF-TESTS

↑ reflection & linking of Info
↑ realism about uncertainty
Highlights misconceptions
Students like it

EXAMS

↑ psychometric reliability
↑ psychometric validity
↓ question numbers
No loss of conventional information

CBM doesn't require special Questions

E.g. You can use past exam Qs to make CBM self-tests
(True/False, Single Best Answer MCQs, EMQs,
Simple text, Numerical, etc.)

- anything with right/wrong marking

CBM exams yield conventional data in parallel

- Accuracy, Item Response data
- a help for standard setting

Knowledge crucially depends on certainty

- A lucky guess is not knowledge
- A firm misconception is far worse than acknowledged ignorance

So why do we usually mark students as if these things weren't true?

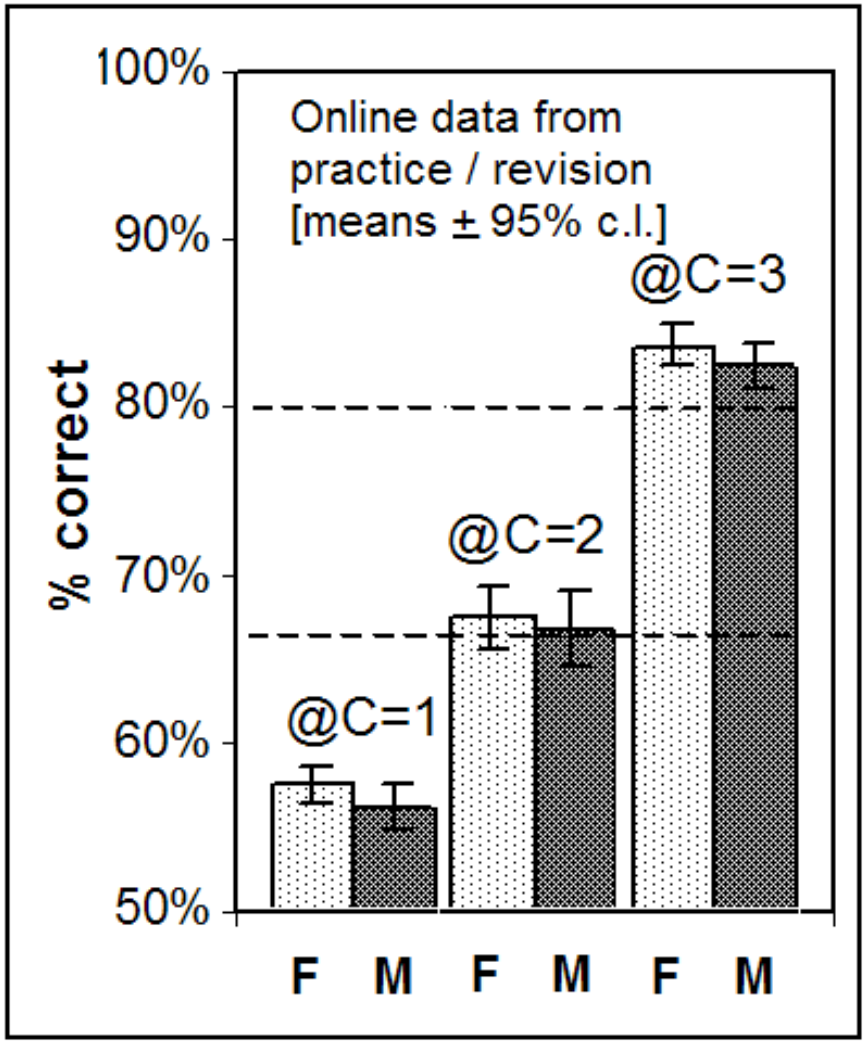
positive marks

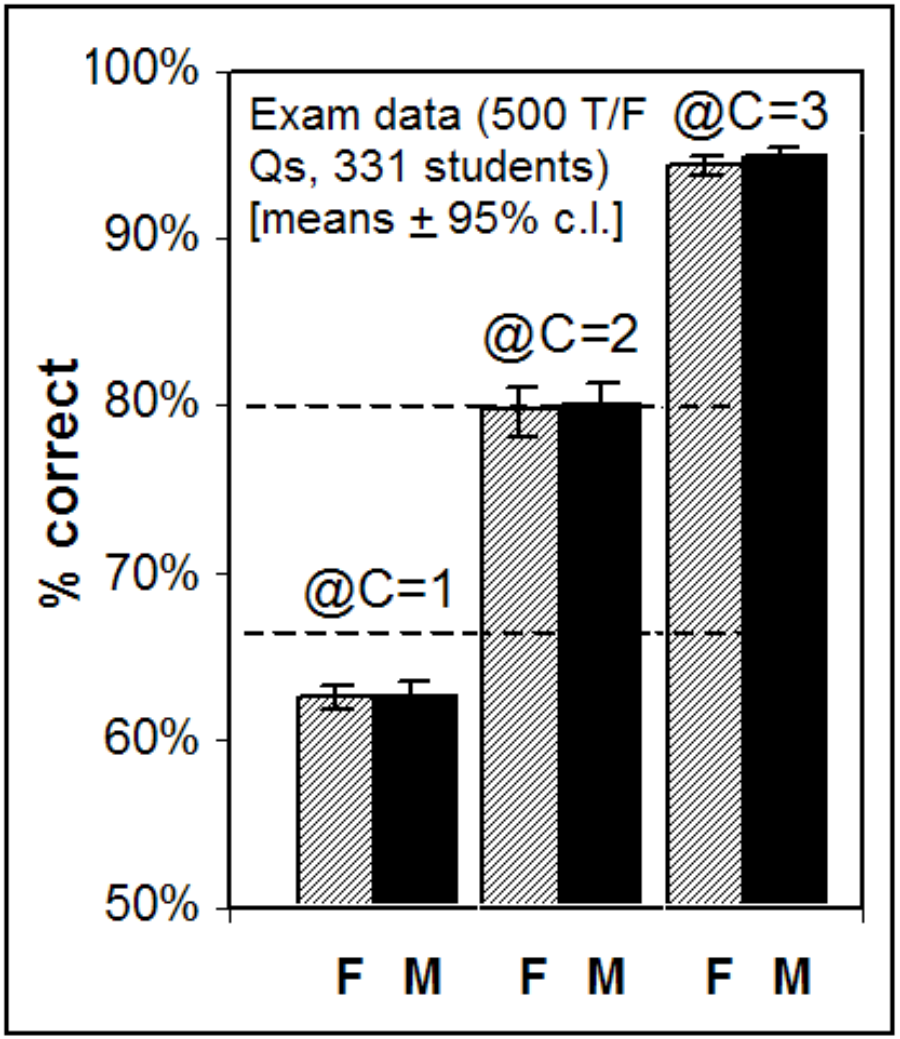
knowledge = zero

negative marks

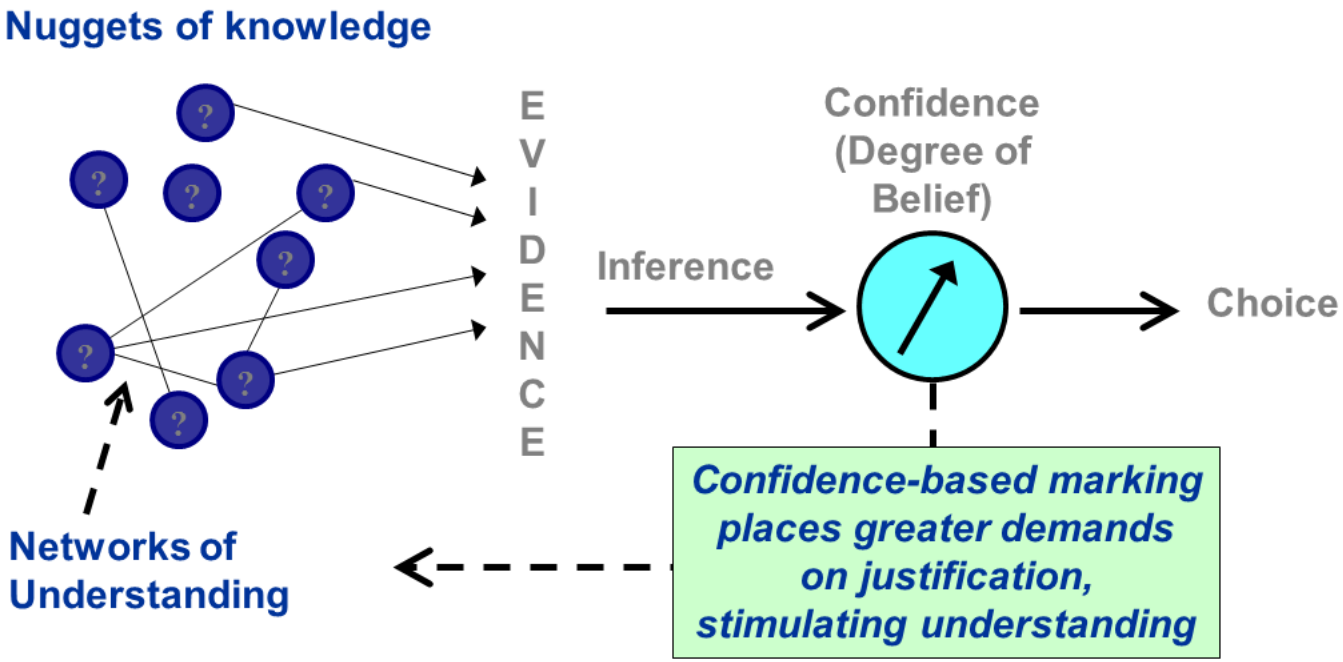
✓ knowledge	*Decreasing* certainty about what is true.
✓ uncertainty	
? don't know	*Increasing* certainty about something false.
× misconception	
✘ delusion	Increasing "ignorance"

How well do students discriminate reliability ?





Thinking about justification and uncertainty stimulates understanding



To understand = to link correctly the facts that bear on an issue.