Adelaide Webinar: 6th April 2011
Transforming Assessment

UCL

Self-test exercises as learning tools, with Certainty Based Marking (CBM)

Tony Gardner-Medwin - Physiology (NPP), UCL

- **The why and the how of Self-Tests**
- **What can CBM add?**
- **Inside or outside a VLE?**
- **Wiki-editing of self-tests**
- **CBM in exam assessment**

Publications, Software, try-out, contact , etc:
www.ucl.ac.uk/LAPT

WHAT ARE WE AS TEACHERS TRYING TO DO?

- Get students to think: to link and relate knowledge
- Use our own (teachers') time to best effect
- Enable students to take charge of their learning
- Discourage superficial (rote) learning
- Take advantage wherever technology has good features

STUDENT ACTIVITIES	TEACHER INVOLVEMENT	THINKING PRESSURE
Listening	high	medium
Practical sessions	high	mixed
Presenting knowledge	high	high
Reading	low	medium
Testing / challenging	low	high
Exams	high	(high)

Assessment – a word to be wary of !

Assess ... from Old French *assesser*, based on Latin *assidere* 'sit by' (in medieval Latin 'levy tax'), from *ad-* 'to, at' + *sedere* 'sit'. [Oxford]

For many teachers, assessment = exams

purpose = grading, filtering, motivating

For learning, the value of assessment (in-course A, peer-A, self-A) is in:

challenge, reflection, feedback

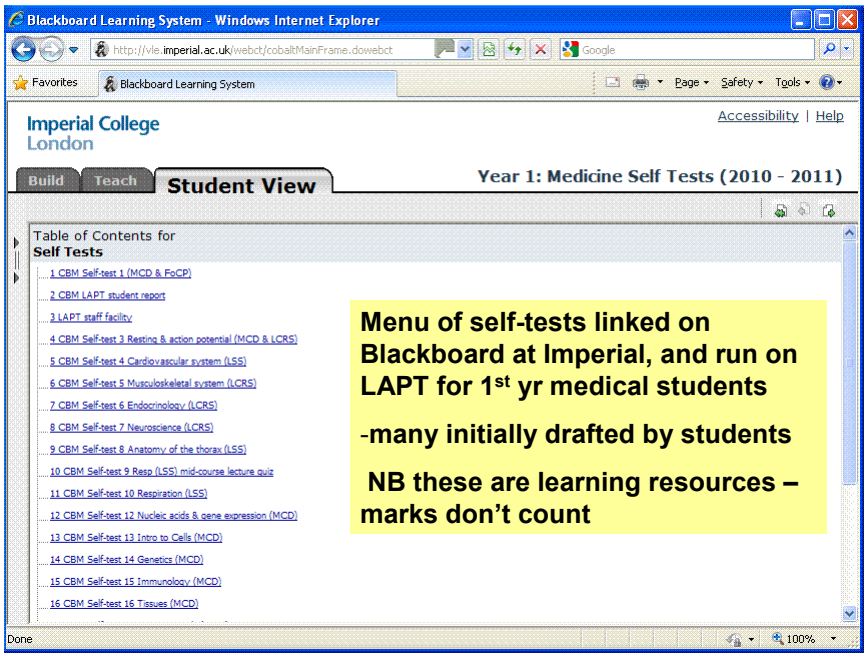
How do we learn to sing, play tennis, or do algebra?

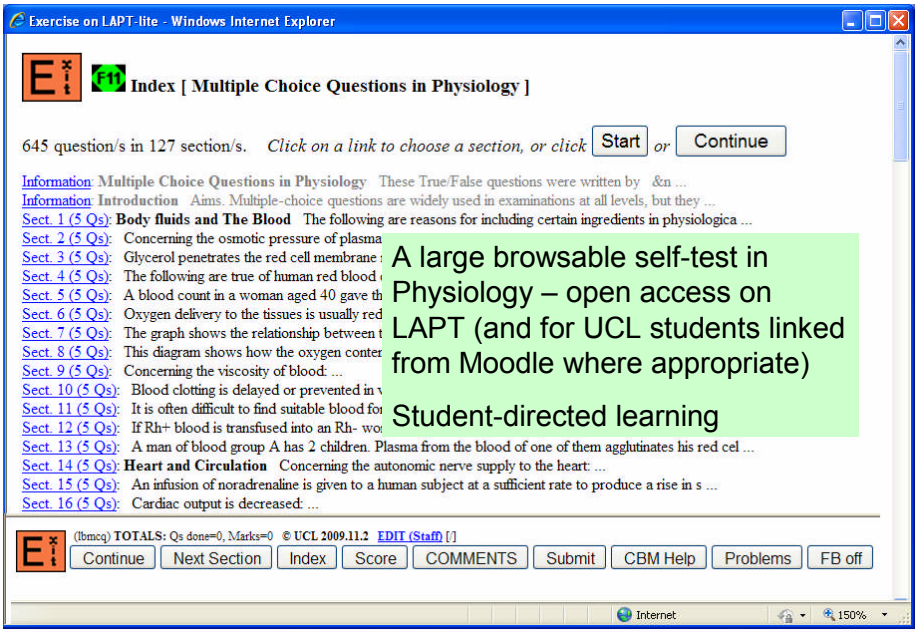
.... by practice, challenge, and thinking

.... by having someone 'sit by' and give feedback

.... but mostly the first

I like to talk about 'self-tests', 'practice' or 'exercises' – not 'exams', 'tests', 'assessments' or 'quizzes' !


Blackboard Learning System - Windows Internet Explorer
Imperial College London
Accessibility | Help
Build
Teach
Student View
Year 1: Medicine Self Tests (2010 - 2011)
Table of Contents for
Self Tests
1 CBM Self-test 1 (MCD & FoCP)
2 CBM LAPT student report
3 LAPT staff facility
4 CBM Self-test 3 Resting & action potential (MCD & LCRS)
5 CBM Self-test 4 Cardiovascular system (LSS)
6 CBM Self-test 5 Musculoskeletal system (LCRS)
7 CBM Self-test 6 Endocrinology (LCRS)
8 CBM Self-test 7 Neuroscience (LCRS)
9 CBM Self-test 8 Anatomy of the thorax (LSS)
10 CBM Self-test 9 Resp (LSS) mid-course lecture quiz
11 CBM Self-test 10 Respiration (LSS)
12 CBM Self-test 12 Nucleic acids & gene expression (MCD)
13 CBM Self-test 13 Intro to Cells (MCD)
14 CBM Self-test 14 Genetics (MCD)
15 CBM Self-test 15 Immunology (MCD)
16 CBM Self-test 16 Tissues (MCD)
Menu of self-tests linked on Blackboard at Imperial, and run on LAPT for 1st yr medical students
-many initially drafted by students
NB these are learning resources – marks don't count


Exercise on LAPT-lite - Windows Internet Explorer
Index [Multiple Choice Questions in Physiology]
645 question/s in 127 section/s. Click on a link to choose a section, or click Start or Continue
Information: Multiple Choice Questions in Physiology These True/False questions were written by &n ...
Information: Introduction Aims. Multiple-choice questions are widely used in examinations at all levels, but they ...
Sect. 1 (5 Qs): Body fluids and The Blood The following are reasons for including certain ingredients in physiologica ...
Sect. 2 (5 Qs): Concerning the osmotic pressure of plasma
Sect. 3 (5 Qs): Glycerol penetrates the red cell membrane
Sect. 4 (5 Qs): The following are true of human red blood
Sect. 5 (5 Qs): A blood count in a woman aged 40 gave th
Sect. 6 (5 Qs): Oxygen delivery to the tissues is usually red
Sect. 7 (5 Qs): The graph shows the relationship between t
Sect. 8 (5 Qs): This diagram shows how the oxygen conter
Sect. 9 (5 Qs): Concerning the viscosity of blood: ...
Sect. 10 (5 Qs): Blood clotting is delayed or prevented in v
Sect. 11 (5 Qs): It is often difficult to find suitable blood for
Sect. 12 (5 Qs): If Rh+ blood is transfused into an Rh- wo
Sect. 13 (5 Qs): A man of blood group A has 2 children. Plasma from the blood of one of them agglutinates his red cel ...
Sect. 14 (5 Qs): Heart and Circulation Concerning the autonomic nerve supply to the heart: ...
Sect. 15 (5 Qs): An infusion of noradrenaline is given to a human subject at a sufficient rate to produce a rise in s ...
Sect. 16 (5 Qs): Cardiac output is decreased: ...
(lbmcq) TOTALS: Qs done=0, Marks=0 © UCL 2009.11.2 EDIT (Staff) [/]
Continue
Next Section
Index
Score
COMMENTS
Submit
CBM Help
Problems
FB off
A large browsable self-test in Physiology – open access on LAPT (and for UCL students linked from Moodle where appropriate)
Student-directed learning

What are desirable features in SELF-TESTS?

- Immediate feedback for each Q
- A stimulating / didactic sequences of Qs
  - mix easy & difficult Qs : for engagement, reward, realism
  - include classic misconceptions – learn through mistakes
  - make chains of Qs : lead through the logic of a topic
- Explanations should widen an issue into other contexts
  - prompt the bringing together of different kinds of knowledge
- Clear navigation – students should be choosing what to do
- Allow use of study materials – tests shouldn't be time-limited
- Encourage comments & dialogue
  - linked to specific quiz/question contexts
- Encourage working in pairs (or more)

- Certainty Based Marking (CBM)

Certainty-Based Marking (CBM)

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

Probability Correct? **<67%** **>80%**

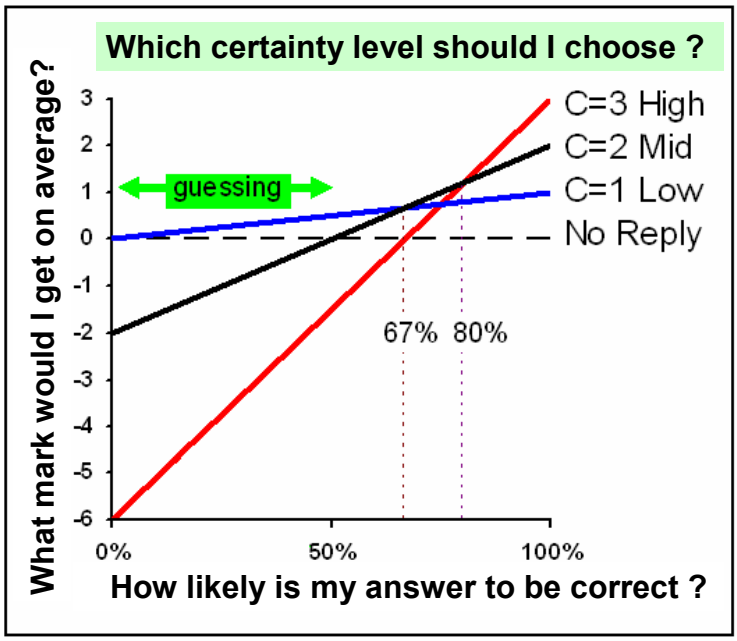
To do well, you need to do more than just jump at an answer:

- can you justify being sure (C=3), by finding links to other knowledge?
- are there uncertainties : reasons for reservation (C=1)?

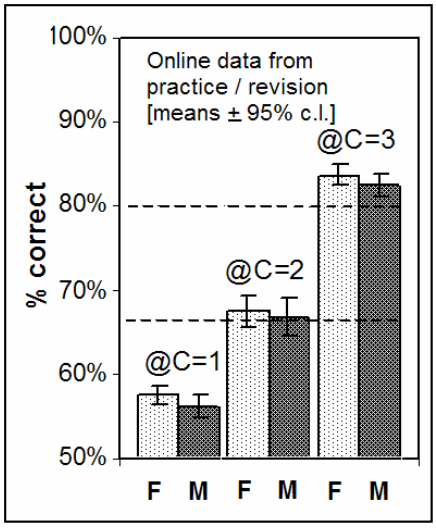
Misconceptions (confident errors) are highlighted;

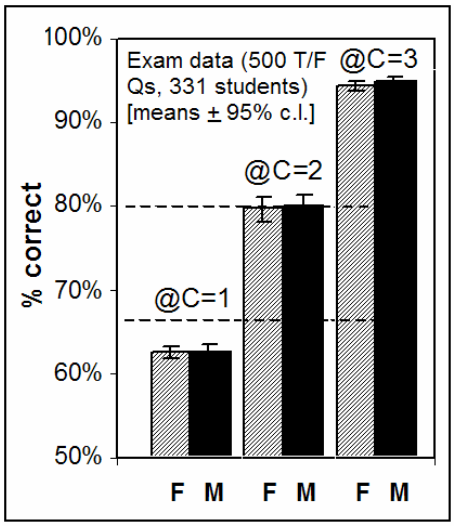
- a double negative mark (-6) hurts!

CBM engages students more (and enhances assessment reliability)



How well do students discriminate reliability ?





LAPT (www.ucl.ac.uk/LAPT)

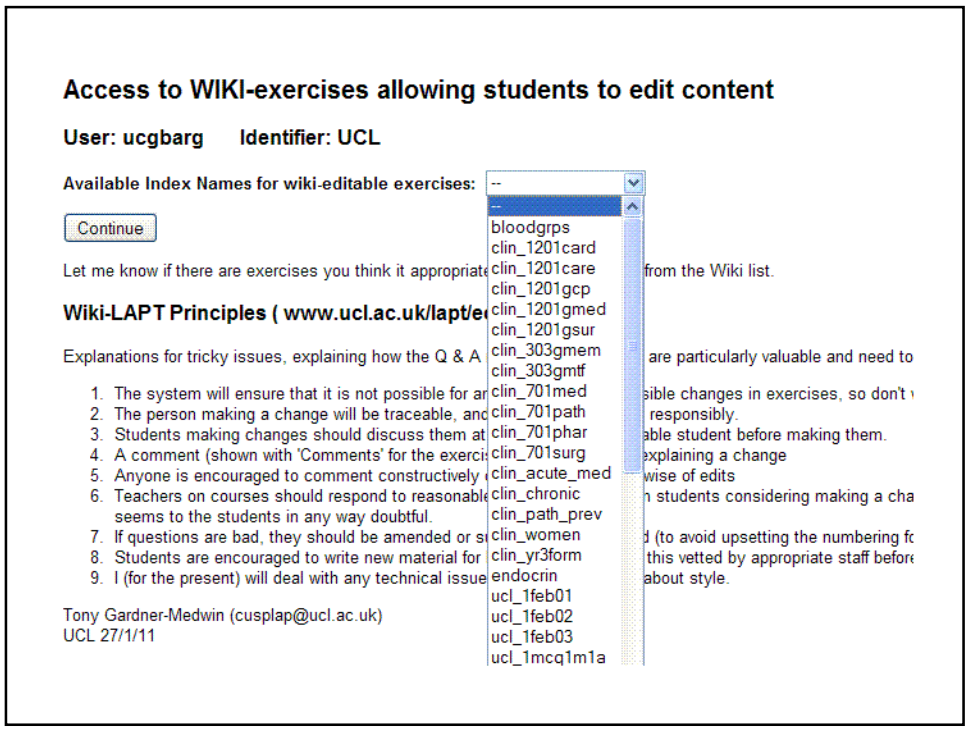
has been developed (since '94) specifically for self-tests in science, medicine & basic maths, and to incorporate CBM efficiently

Can I use self-tests in my VLE / LMS?

Yes, of course. But LAPT has features you may struggle to include:

1. CBM [CBM code can go in Moodle, but Moodle is always changing]
2. Immediate feedback independent of server-interaction [with .JS]
3. Marks relative to the Qs a student chooses, not the entire file
4. Comment & discussion facilities for individual questions
5. Grouped questions, with common stems (cf. common medical practice, Questionmark)
6. Editing of exercises as whole simple files, not individual questions
7. Wiki-editing facilities – students can update/ correct/ explain points

I prefer authenticated links to LAPT from within a VLE, with students able to see their session reports on LAPT



Is certainty just 'meta-cognition' or is it fundamental to knowledge?

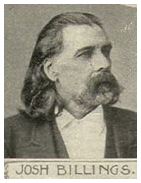
Ordinary words we use to describe Knowledge :

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

- *Knowledge is a function of certainty (confidence, degree of belief)*
- *There are states a lot worse than acknowledged ignorance*

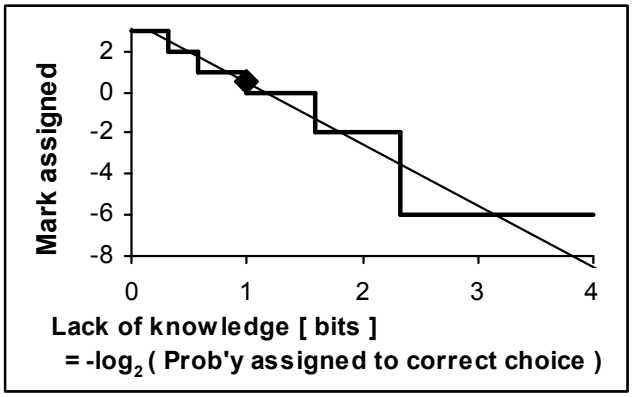
"***It's not ignorance does so much damage - it's knowin' so derned much that ain't so.***"

attrib J. Billings



Should I use CBM then in exams?

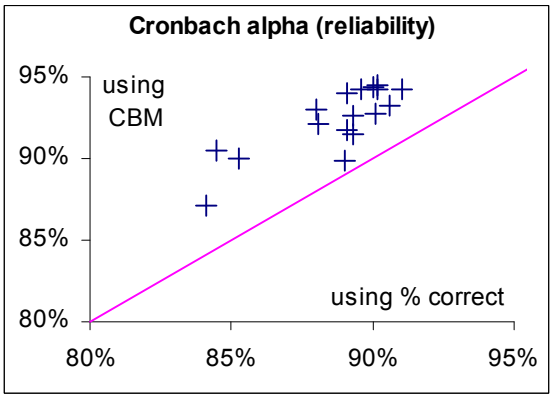
CBM does quite closely follow the ideal measure of ignorance



The student loses about 3 marks per 'bit' of ignorance - up to a maximum of 3 bits

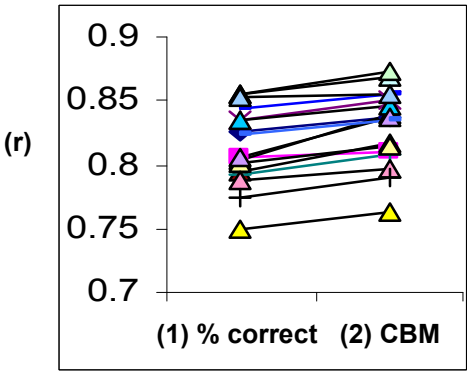
... and CBM increases the reliability of exam data

'Reliability' indicates how well a score measures something about the student's ability, as opposed to 'luck' or chance.



To achieve these increases using only % correct would have required on average 58% more questions.

... and CBM scores are better predictors of how many correct answers a student will give on a different set of questions



Correlation coefficients (r) showing how well the % correct on odd numbered Qs was correlated with (1) % correct and (2) CBM scores on even numbered Qs, from 17 exams (250-300 Qs each), >300 students.

Are there problems in using CBM in exams?

Prior practice with CBM is essential (lots of self-tests!)

Adjustment for students with poor calibration may be fair

Standard setting requires care, with a new approach

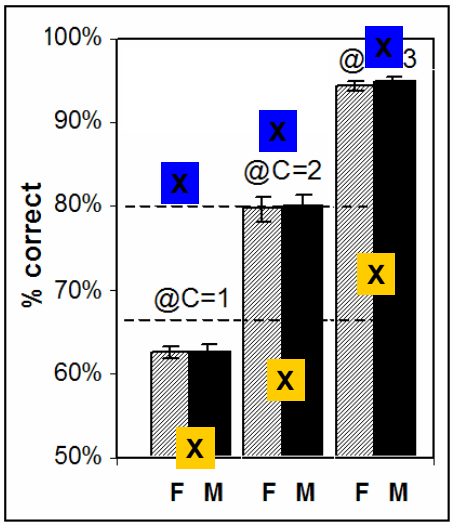
Exam boards do tend to be wary of innovation

BUT:

CBM exams yield extra data not different data, so the percent correct scores are always there, for comparison with traditional techniques

How you mark exams motivates how students will study in self-tests, for which CBM is even more clearly of value

Adjustment for poor calibration



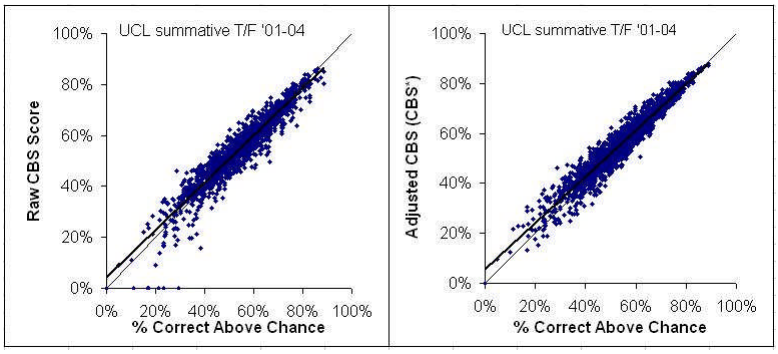
Significantly underestimate the reliability of their answers: e.g. 41 students (14%)

Significantly overestimate the reliability of their answers : e.g. 2 students (1%)

Maybe one shouldn't penalise such students

.... you can adjust their scores by changing e.g. C=2 to C=3 or C=3 to C=2

Adjustment of Scores can take account of broad mis-calibration of confidence

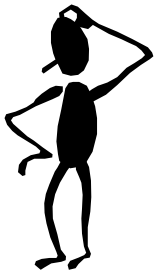


So why doesn't everybody already use CBM ? - a puzzle

- Enthusiasm was exhausted before the age of 'online'
- Some CBM methods were complex, opaque or non-motivating
- Reluctance to treat certainty as integral to knowledge
- Mistaken worries about 'personality bias'
- Under-rating of self-tests & practice as learning tools
- Worry that CBM would need new questions
- Worry that CBM would upset standard-setting
- Inertia and vested interests

www.ucl.ac.uk/LAPT
a.gardner-medwin@ucl.ac.uk

UCL



www.ucl.ac.uk/LAPT

a.gardner-medwin@ucl.ac.uk

How should one handle students with poor calibration?

Significantly overconfident in exam: **2 students (1%)**

e.g. 50% correct @C=1, 59%@C=2, 73%@C=3

Significantly underconfident in exam: **41 students (14%)**

e.g. 83% correct @C=1, 89%@C=2, 99%@C=3

Maybe one shouldn't penalise such students

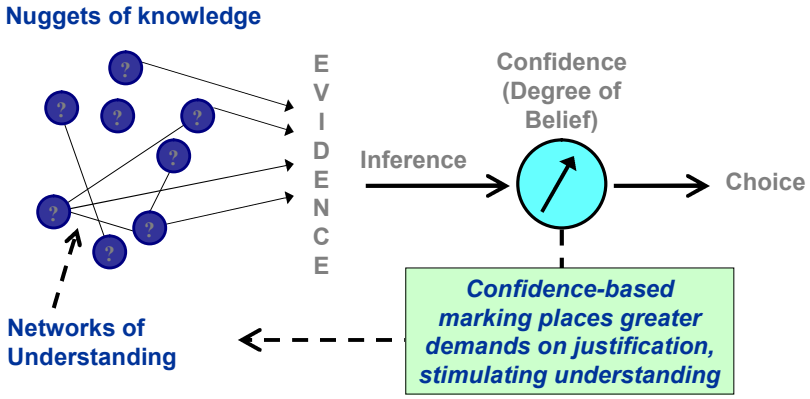
Adjusted confidence-based score:

Mark the set of answers at each C level as if they were entered at the C level that gives the highest score**.

mean benefit = 1.5% ± 2.1% (median 0.6%)

** (first combining sets if %correct is not in ascending order)

Thinking about uncertainty and justification stimulates understanding



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