Certainty Based Marking: Why, How & When?
Tony Gardner-Medwin – UCL – ugbarg@ucl.ac.uk

• 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? Ideas, Reservations?
• My motivations? What is knowledge?
• CBM, proper marking schemes, self-tests vs assessments
• CBM: performance in self-tests & exams, CB ‘bonus’ concept
• Implementation: LAPT, MOODLE, private offline self-test modules

Publications, software, try-out, contact, etc:
www.ucl.ac.uk/lapt www.TMedwin.net/cbm (new modules)

Irrationality of fixed negative marking

Students who have the insight to identify which of their answers are unreliable may omit these (perhaps following misguided instruction)
The result is they will on average do worse than students without such insight. This is quite improper and such schemes should be illegal.

IDEAS & RESERVATIONS

A lucky guess is not knowledge
NB a guess is usually informed by some knowledge
A confident misconception is worse than ignorance
It certainly can inhibit learning and can be dangerous.
But NB misconceptions may reflect genuine knowledge about something related, e.g. “Australia’s capital is Auckland”.
We generally ignore these things
NB some people think (incorrectly!) that negative marking helps, by discouraging guessing - or they scale scores so guesses will on average give zero marks.
Why do we ignore them?
 Desire to encourage answers based on partial knowledge?
Dislike of negative marking?
Averaging is supposed to get round the problems?
Concern that certainty/uncertainty are different from knowledge? 
........... or maybe not measurable?

MY MOTIVATIONS

Teachers
Free them to do what they do best: stimulating Interest, Creativity, Understanding, Relationships
Use IT to help them, not replace them!
Use meta-information, as in face-to-face interaction.
Reward acknowledgement of uncertainty
Stimulate thinking, not rote learning
Encourage control of their own learning
Learning through mistakes, misconceptions
Self-challenges in private, like music/sports practice

Online data from practice / revision
[means ± 95% c.i.]

How well do students discriminate reliability?

<table>
<thead>
<tr>
<th>Degree of Certainty</th>
<th>C=1 (low)</th>
<th>C=2 (mid)</th>
<th>C=3 (high)</th>
<th>No Reply</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mark if correct</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>Penalty if wrong</td>
<td>0</td>
<td>-2</td>
<td>-6</td>
<td>0</td>
</tr>
<tr>
<td>Probability Correct</td>
<td>&lt;67%</td>
<td>67-80%</td>
<td>&gt;80%</td>
<td>-</td>
</tr>
</tbody>
</table>

Student perspective:
• Always motivated to be honest
• Rewarded for identifying weaknesses
• Rewarded for sound justifications
• Encouraged to reflect & link info
• Misconceptions highlighted
• Simple and transparent scheme
• Perceive it as realistic & fair

Staff perspective:
• Doesn’t require new or different Qs
• Enhanced feedback about content
• Enhanced reliability & validity in exams
• Better student learning experience
What is knowledge?
- Knowledge = justified true belief
- Certainty = degree of belief
- Justification requires understanding

Decreasing confidence in what is true,
Increasing confidence in what is false

What is understanding?
- To understand = to link correctly the facts that bear on an issue.

Certainty-Based Marking places greater demands on justification, thereby stimulating understanding

Nuggets of knowledge
Inference
Choice
Certainty (Degree of Belief)

Networks of understanding

CBM Self-tests: what the marks tell you

Very good, but may have repeated self-tests excessively
Good insight into which knowledge is reliable
Underestimates, knowledge, or not serious about CBM
Knows quite a bit, but doesn’t know where shaky

Misconceptions or lack of awareness of ignorance

CBM mark if you use the same C all the time

Little knowledge but knows what he doesn’t know

Performance in January Formative: first on-paper test in Med Sch

Students who did Self-tests

Students who did NOT do Self-tests

Students who did NOT do Self-tests are about twice as likely to fail as students who did Self tests.
Pattern similar every year: Use is a good predictor of Formative performance

Results for Jan2012

N.A. Curtin, Imperial College

Presentation of CBM marks: introduction of a CB 'bonus' concept

CBM FEEDBACK EXAMPLES (on LAPT: Imperial Self-Test)

Little knowledge but knows what 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

Music practice

Model for Learning

Sports Practice

Cooperative Challenging
You learn from mistakes
You mark your boundaries
You push them Out of view of your teacher
Fun

The elements:
Thinking
Challenging
Practising
Correcting
Floundering
Selecting
Discussing
Enjoying
CBM in Exams

- Standard setters get conventional accuracy (% correct) as well as CBM
- For the same accuracy, students gain if they correctly identify strengths and weaknesses
- CBM is a more soundly based measure of ignorance or knowledge
- CBM yields exam data with greater statistical reliability
- CBM is better than accuracy for predicting the accuracy on a separate set of Qs

Data from 1000 random splits of 17 exams (250-300 T/F Qs) into equal subsets. Correlations between student rank order on each set, based on Accuracy or CBM:

- ↑ of reliability with CBM was equivalent to a 42% ± 7% (sem) ↑ of Q numbers
- ↑ of predictive power for accuracy was equivalent to a 2.2% ± 1.5% (sem) ↑ of Q numbers

CBM enhances reliability and validity of exam scores

CBM IMPLEMENTATION - How?

SELF-TESTS
- LAPT (London Agreed Protocol for Testing) open access self-tests and you can use your own authenticated access and your own self-tests. Contact me.
- You can link to LAPT from your LMS (BlackBoard at Imperial, Moodle at UCL etc.)
- MOODLE: CBM is now well implemented in Moodle 2.6 (Nov 2013) Moodle 1.9 – 2.5 require code patches (TMedwin.net/cbm)
- New Self-Test software at TMedwin.net/cbm will replace LAPT, with installation wholly on your own server, plus downloadable personal self-test modules.

SECURE EXAMS*
- Moodle 2.6
- Optical Mark Reader Sheets (Speedwell)

* LAPT & similar software for self-testing is not recommended for secure testing, because feedback is provided locally, rather than from a server.

SUMMARY - Why? When?

CBM makes Sense!
- Is easily implemented
- Doesn’t require writing special questions
- Always motivates students to identify & acknowledge uncertainty

<table>
<thead>
<tr>
<th>SELF-TESTS</th>
<th>EXAMS</th>
</tr>
</thead>
<tbody>
<tr>
<td>↑ reflection &amp; cross-linking of Info</td>
<td>↑ psychometric reliability</td>
</tr>
<tr>
<td>↑ realism about uncertainty</td>
<td>↑ number of questions required</td>
</tr>
<tr>
<td>Highlights misconceptions</td>
<td>Familiar standard-setting info retained</td>
</tr>
<tr>
<td>Challenge and practice in private</td>
<td>Students understand and value CBM</td>
</tr>
</tbody>
</table>

Contributors to the project, over many years:
- David Bender, Nancy Curtin, Chris Dean, Mike Gahan, Kim Issroff, UCL & Imperial students
- Earlier pioneers of work on confidence assessment & learning:
  - Andrew Ahlgren, Jim Bruno, Robert Elbro, Jack Good, Kate Hever, Darwin Hunt, Dieudonné Leclercq, Emir Shufeld

Contact: Tony Gardner - tmedwin@ucl.ac.uk