

Things I'm not going to do

1. Compare knowledge in one topic with another
- though I shall compare ignorance
2. Give a recipe for deciding if something is true or probable - the 'science of evidence' [Phil Dawid's Leverhulme Project]
3. Discuss relativism - frameworks of belief - whether knowledge comes from science, religion, experiment, argument, etc.

*I assume we start with a defined body of accepted consistent truths (a '**TOPIC**'), and we are interested in whether we can identify and measure how much knowledge or ignorance someone has about these*

How do we assess knowledge?

Approaches

Ask for an exposition of K

Ask for inferences made from K

Ask Qs and count correct answers ...

Problems

Was it rote-learned?
Would that matter?
Is understanding a part of K?

Are they based on guesses?
How hard are the Qs?

... all of the above

Are there real standards? Do assessors agree?

... "65% is a 2:1" ... 65% of what? ... "If you think it's a 2:1, put 65%"

... is assessment of knowledge just a matter of opinion?

Is there any gold standard - a definition of Knowledge or Ignorance that gives the same answer independent of details of protocol?

Measuring knowledge or ignorance within a topic

1. Identify the topic - a set of true propositions
2.
 - choose a proposition P whose truth or falsity is implied by the topic
 - ask the subject to assign a probability B (degree of Belief) that P is true
 - $C=B$ if P is true; $C=1-B$ if P is false. (C is probability for a Correct inference)
 - add $\log(1/C)$ to a cumulative score of ignorance, I
 - tell the subject whether the proposition P was T or F (*feedback*)
3. Repeat this till you have covered propositions *equivalent* to the topic (i.e. whose truth or falsity entails and is entailed by those in the topic)

The result is **IGNORANCE of the topic**, independent of :-

- a) *the choice of questions,*
- b) *whether they were part of the initial definition of the topic,*
- c) *how similar or repetitive they may be, or their order,*
- d) *how they relate to the way in which the subject learned the topic*

... **PROVIDED** that the subject:

1. holds consistent probabilities, and reports these correctly
2. is able to form inferences from propositions within the topic
3. uses feedback rationally to inform future responses

Any deviation from 1,2,3 will *increase* the ignorance measure (on average)
No other measure is independent of a,b,c,d

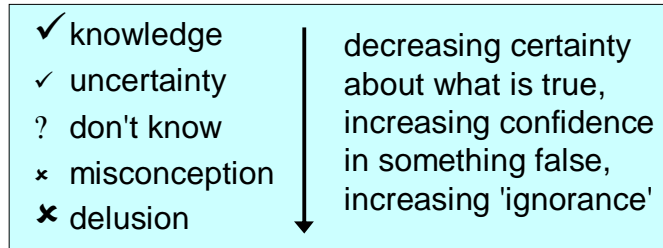
Reservations?

- We can't realistically do this for substantial topics - it would take as long as learning the topics. We must rely on sampling.
- Real people will not be fully consistent or rational, so their ignorance score will be raised on that account, by an amount depending on the protocol

Conclusions from theory?

- Ignorance has a unique measure, if the subject is consistent and rational
- Knowledge is inseparable from understanding. Without understanding, ignorance is ill-defined and potentially unbounded
- Parrots are ignorant, however well they can recite a topic
- Students likewise, if they cannot make inferences within the topic
- Ignorance for a statement is $-\log(\text{probability assigned to its truth value})$
- Ignorance cannot be derived from just correct responses & errors
- Exposition of a topic is an unreliable way to test knowledge, because it may overlook (a) lack of understanding and (b) uncertainty
- Learning technology is well suited to obtaining the required information

Ordinary ways we think about Knowledge



*Knowledge is a function of confidence (certainty, or degree of belief)
The only anchor point is the top: total knowledge or zero ignorance
There are states a lot worse than acknowledged ignorance*

"It ain't what you don't know that gets you into trouble.
It's what you know for sure that just ain't so."

- attrib. to Mark Twain
[Al Gore - *An Inconvenient Truth*]

'An Inconvenient Truth'



***If you haven't seen it,
book now at the Odeon, Panton St.!***

Epistemology?

Knowledge is justified true belief

Truth - OK, we are assessing relative to given truths

Belief - We certainly have degrees of belief - a person may be certain, uncertain, or simply guessing - realistically, a continuum of probabilities

Justification - The trickiest and most important concept, critical to both assessment and learning

[Gettier? - discussion]

NB modern pedagogic focus is on "Assessment for Learning" (e.g. Black, Sadler, Knight). Assessment should be part of the learning process, encouraging students' thinking about justification of answers, tying them in with other knowledge, implications and whether these enhance or go against belief in a conclusion. Assessment must encourage and reward such thinking. Reliable self-assessment is both constructive for learning and a desired outcome of education.

Your degree of belief is determined by your attempts to justify a proposition. Educational assessment should force you to make honest judgments about reliability -- Certainty-Based Marking (CBM)

The UCL (LAPT) Certainty-Based Marking scheme

Certainty Level:	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
P(correct)	<67%	> P <	>80%	-

London Agreed Protocol for Teaching (Of Physiology)



All about LAPT :

www.ucl.ac.uk/lapt

Exercises, Publications, Tools, etc.

How should one choose the best CBM level?

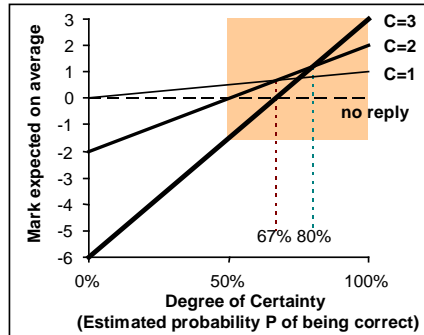
Which line is highest?

C=3 when P(correct) >80%

C=1 when P(correct) <67%

The student is always motivated to be honest about low or high certainty, to get the best score.

[a proper scoring rule - P.Dawid]



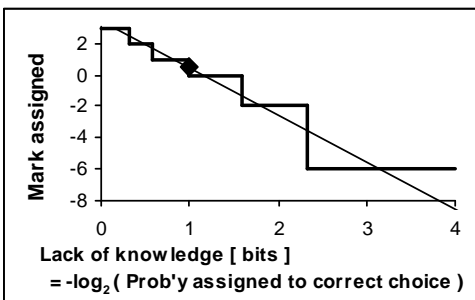
NB the student gains:

EITHER by finding justification for high confidence

OR by seeing reasons for reservation about an answer

If two students give the same answers, the one who distinguishes reliable from uncertain answers will do better.

CBM quite closely follows the ideal ignorance measure



The student loses about 3 marks per 'bit' of ignorance

- up to a maximum of 3 bits

Example Qs

TRIAL EXERCISES:

Knowledge:

UK speed limits for a car (no trailer) on the following roads (not otherwise signposted) are? ...
A= 30 mph B= 40 mph C=50 mph D= 60 mph E= 70 mph

	Confidence
1.→ Dual carriage-way	1 <A> <C> <D> <E> <1> <2> <3>
2.→ Motorway	2 <A> <C> <D> <E> <1> <2> <3>
3.→ Ordinary country road (2-way)	3 <A> <C> <D> <E> <1> <2> <3>
4.→ Single lane country road with passing places	4 <A> <C> <D> <E> <1> <2> <3>

Numeracy:

5. Your antique clock rose by 80% in value in 10 years then fell by 40% in the next 10.
By what % did it rise in value overall, in the 20 years?

Confidence

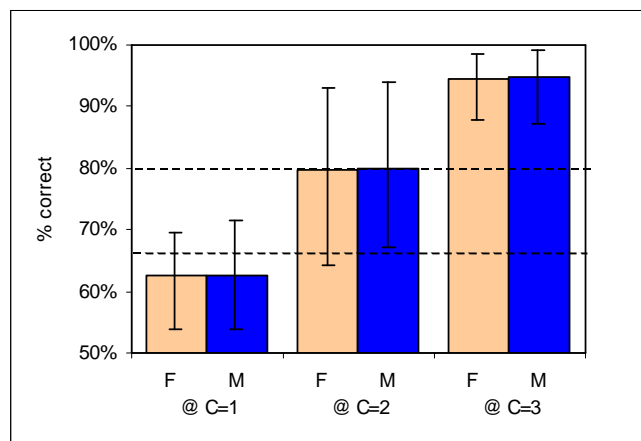
<1> <2> <3>

Logic:

6. Every branchiopod is a crustacean and every crustacean is an arthropod. No insect is a crustacean.
Given this, which of the following must be true? (Mark these T and the others F)

	Confidence
A.→Every branchiopod is an arthropod	A <T> <F> <1> <2> <3>
B.→No insect is an arthropod	B <T> <F> <1> <2> <3>
C.→No branchiopod is an insect	C <T> <F> <1> <2> <3>
D.→Some crustaceans are insects	D <T> <F> <1> <2> <3>

How well do the students discriminate reliability?



Bars show range including 90% of students

[331 students, 500 T/F exam questions, 2001]

Students adopt certainty-based marking very easily

Principles that students seem readily to understand :-

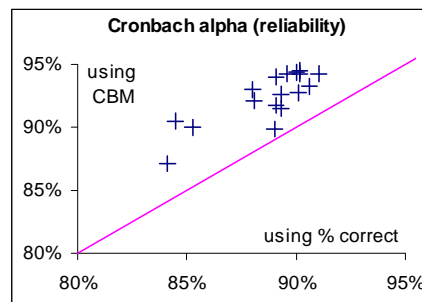
- You must know the reliability of each bit of your knowledge to use it
- confident errors are serious, requiring careful attention to explanations
- expressing uncertainty when you are uncertain is a **good thing**
- confidence is not just personality, it is being able to justify what you say
- if you are over- or under-confident, you should recalibrate with practice
- reflection integrates knowledge and understanding, and should be routine

In evaluation surveys, a majority of students have always said they like CBM, finding it useful and fair, and recently they voted 52%: 30% that it should be retained in Year 1&2 medical exams.

CBM in Exams with True/False Questions

17 medical exams, 250-300 questions in each, ca. 300-350 students

The standard measure of exam reliability, as a measure of student ability rather than chance factors, is called Cronbach Alpha.



Reliability increases if you use more exam questions.

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

Extended Matching Questions (EMQs)

- a new fashion in medical assessment

CBM Trial (Feb '06) with EMQs

- Weak students over-estimated the reliability of their EMQ answers.
- Adjustment can be made to compensate, but ...
- this occurs much less with T/F Qs, and therefore CBM may be especially valuable with EMQs.
- *The Medical School has puzzlingly discontinued CBM trials with EMQs.*

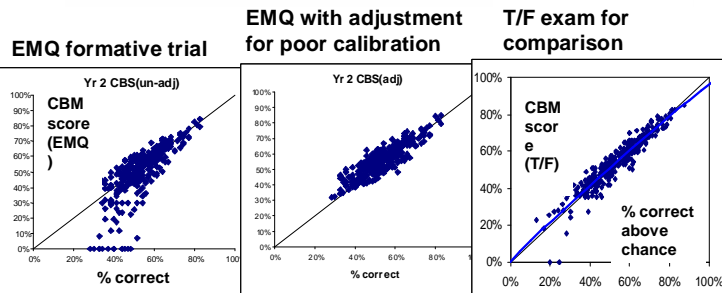
Alcohol intoxication	Meningococcal septicaemia
Complete heart block	X Pulmonary embolus
Epilepsy	Supraventricular tachycardia
Gram negative septicaemia	Vaso vagal faint
Left ventricular failure	
Mechanical fall	

X Mark = -2 (conf=2)

For each of the scenarios below select the most likely diagnosis from the list above:

Qu. 1: An 80 year old man presents with collapse associated with severe shortness of breath. He has crackles throughout both lung fields.

Model Answer: Left ventricular failure



CBM - main current usage

Learning online or downloaded: self-assessment (> million/year)

- UCL, Imperial + students at >30 universities
- Follow-up exercises for lectures, practicals
- Revision, with past exams
- Student-written exercises (UCL & Imperial)

Formal online tests (using link to WebCT or VLE)

- Maths & key skills tests at UCL
- Formative module tests at Imperial
- Student teacher audit at Winchester College

Paper (OMR) tests

- Formative module tests at UCL & Imperial
- Yr 1,2 exams at UCL

Lecture/ seminar context

- Junior Doctor sessions at Imperial [Dr. Sara Marshall]
"Are you prepared to act on your answer?"

Lessons from experience with CBM

Practice is needed before use in exams

- not really a problem, since its objective is to encourage better thinking

Exams should re-use questions from an open database only very sparingly

- CBM places a premium on answers that the student has good reason to believe are correct - having seen the Q&A before is a good reason!

Students can lose out through excessive self-confidence or risk-aversion

- these are unhealthy traits that are moderated by practice with CBM
- adjustment can be made in exams to compensate for poor calibration**

With Best-Option Qs, students often think answers more reliable than they are

- further research required
- students may benefit from confronting reality by using CBM

Standard setting

- the CBM mark range is unfamiliar, but scaling aligns it with the familiar**

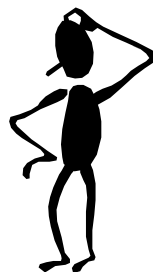
**** see Publications**

Positive Features :

- students like CBM, consider it helps them study and is more fair
- they have voted at UCL to retain it in Yr 1,2 exams
- it is more reliable and valid than conventional marking in exams
- it is more closely related to what we mean by 'knowledge'

We fail if we mark a lucky guess as if it were knowledge.

We fail if we mark delusion as no worse than ignorance.



www.ucl.ac.uk/lapt

Lunch Hour Lecture 21 Nov 2006, Introductory Remarks

- How does the brain handle Knowledge and Uncertainty?
- Years ago I shifted priority from this question to how university staff handle knowledge and uncertainty.
- I felt this vastly more important.
- I thought, foolishly, that people would rapidly see sense in obviously good ideas that have been around but dormant for decades.
- Needless to say, most university staff behaved rather like hippopotamuses, neither budging nor arguing.
- I'm still interested in the brain of course, but there seem to be masses of people working on Cognitive Mechanisms these days, not least
at the Gatsby Unit in Queen Square - where I transferred research funds
- There are few paying attention to what I'm going to talk about today.

I've shifted to my pension after a 40 year slog, probably much to the financial relief of the Provost. I don't call it retirement because I have never intended less work, merely freedom to do what I think most important. Remarkably, I lose only 15% of net income, a small price for freedom!

[EAR] If I can persuade audiences like you to take seriously the science of knowledge measurement, then I will be able to spend more of my time on ear physiology - the Organ of Corti - one of the most beautiful pieces of physiological machinery in the body. I like to put this up because it gives me a twinge of guilt that, even without a job, I still don't have enough time to work on it properly.

[SLIDE ENCYC] This is from the 1751 French Encyclopedie of Diderot and D'Alembert, one of the great gems of the enlightenment. For fun, I've ringed some of the things I touch on today. If I did what UCL would like me to do, I doubtless would stay beavering in a little backwater off the screen called Physiology. But I want to draw attention to the structure at the top here.

Knowledge embraces understanding (French 'entendement' - perhaps too subtle to translate properly to English), memory, reason and imagination. Current thinking often separates 'knowledge' from 'understanding'. I shall argue that this is quite wrong. Understanding and reason are inseparable from knowledge. Teachers must battle against rote learning as the way to pass exams. If we have a parrot that can recite the rules of tennis, the parrot does not know the rules of tennis. Unfortunately the way we teach and assess our students all too easily encourages parrot- learning.

If I have a criticism of this map, it is that 'imagination' includes no mention of science - where models, analogies and hypotheses are of the essence.

But let me first identify some things that I'm NOT going to try to do

There was time for one question after the lecture, which I unfairly ducked on the excuse that I had explicitly said in the lecture that I wouldn't talk about it ! Here is a constructed dialogue based on that question.

FD: My question was based in the quote attributed to Mark Twain: '*faith is believing in something you know is not true*'

TGM: '*believing in something you know is not true*' seems to me a straight contradiction, and would surely be rejected by anyone holding something they called a faith. I have faith in rational argument, which I could only justify in a circular manner, but I certainly don't know it to be unmerited. If my faith in it broke down I would be totally at sea - not feeling at all relieved of an untruth!

What could be behind Mark Twain's remark? Change it to: '*Faith is certain belief in something that you know you can't prove*' and you have quite a reasonable definition of faith. It doesn't exclude your having evidence for the belief. But you know you can't provide proof to win round a sceptic. One of the commonest forms of evidence for religious beliefs is the fact that someone for whom you have respect and trust claims it to be true. This is indeed all the evidence that supports many of my scientific beliefs, though I don't claim to hold them with complete certainty.

Where the notion of certain belief ($P=1$) comes from in religion or science I don't understand, and it seems to be the cause of a lot of trouble. If only we thought in terms of odds ($P/(1-P)$), or log-odds, perhaps people would be less inclined to claim the extremes: "*Infinitely certain*" is a lot more risky and implausible sounding than "*completely certain*". This is a serious point, made first I think by Turing and Michie and Good many decades ago, because defining belief as log-odds (extending from minus infinity to plus infinity) rather than as probability (from 0 to 1) gives the scale on which weights of independent evidence (log likelihood ratios) shift us linearly, whatever our prior belief. It encouraged Turing (I think it was) to propose the unit of evidence and belief as the Bel (10 dB or 3.3 bits), thereby boosting the legacy of Alexander Graham Bell.

FD: Does someone that can remember entire passages in the bible, or even explain in detail the mysteries of the catholic faith, have knowledge? There is good memory there and plenty of imagination, but reason?

TGM: Obviously there is a distinction between knowing what's in the bible and believing, or being able to understand and reason from its propositions. My position in the lecture was that you are ignorant of a topic unless you can make rationally consistent inferences from propositions within it. The ability simply to recite them is ignorance, parrot-fashion. If there are rational inferences to be made from tenets of religious faiths (yielding coherent '*topics*' for which lack of knowledge can be well defined) then I can't generally discern them. I simply don't understand statements like "Christ ascended into heaven" as meaningful, capable of being true or false, or a basis for any kind of inference.

If you have a formalism that makes these things meaningful and consistent - with say heaven as a hypothetical land in the sky - then you could have a well-defined measure of lack of knowledge relative to that formalism. A small ignorance score could be a good basis for acceptance into the faith community, but it wouldn't indicate that the formalism had any relation to reality. That's one of the reasons I say you can measure ignorance in a well-defined manner, relative to a coherent set of propositions, but that you cannot have any absolute measure of knowledge.

FD: Can we have knowledge within a belief system? i.e. knowledge without evidence? is faith knowledge? is there understanding within faith?

TGM: Maybe I've covered this. A curiosity is that if you claim certain knowledge within a faith, then later deny it, this is tantamount (in my analysis) to branding your friends (and your former self) infinitely ignorant, perhaps consequently worthless. Maybe this is why people have been prepared to die rather than deny their faith. An artefact of the way we talk about probability! In normal life one does not operate with certain beliefs, since experience warns that neither one's senses, one's reason nor one's teachers and associates are totally reliable.