

Next steps for assessment

Cambridge Schools Conference
December 2015

Gordon Stobart
Emeritus Professor of Education,
Institute of Education,
University College London

The big picture

- We have visions of the future – but what steps must we take to get there?
- Assessment is a social process – does society want change?
- Competing agendas – 21st century learning vs accountability /selection
- Steps towards fit-for-the-future assessment
- Classroom assessment can lead the way – we need to do more than mimic external examinations.

Assessment as a social process

- Examinations are shaped by broader social forces:
 - What is the purpose?
 - Selection by merit - 1000+ years of Chinese Imperial selection exams
 - Competence and certification - Exams as means of 'closure and control' in professions
 - Setting and raising standards
 - University examinations in Oxford and Cambridge
 - Local examinations for middle-class schools
 - Payment by results.

Assessment as a social process

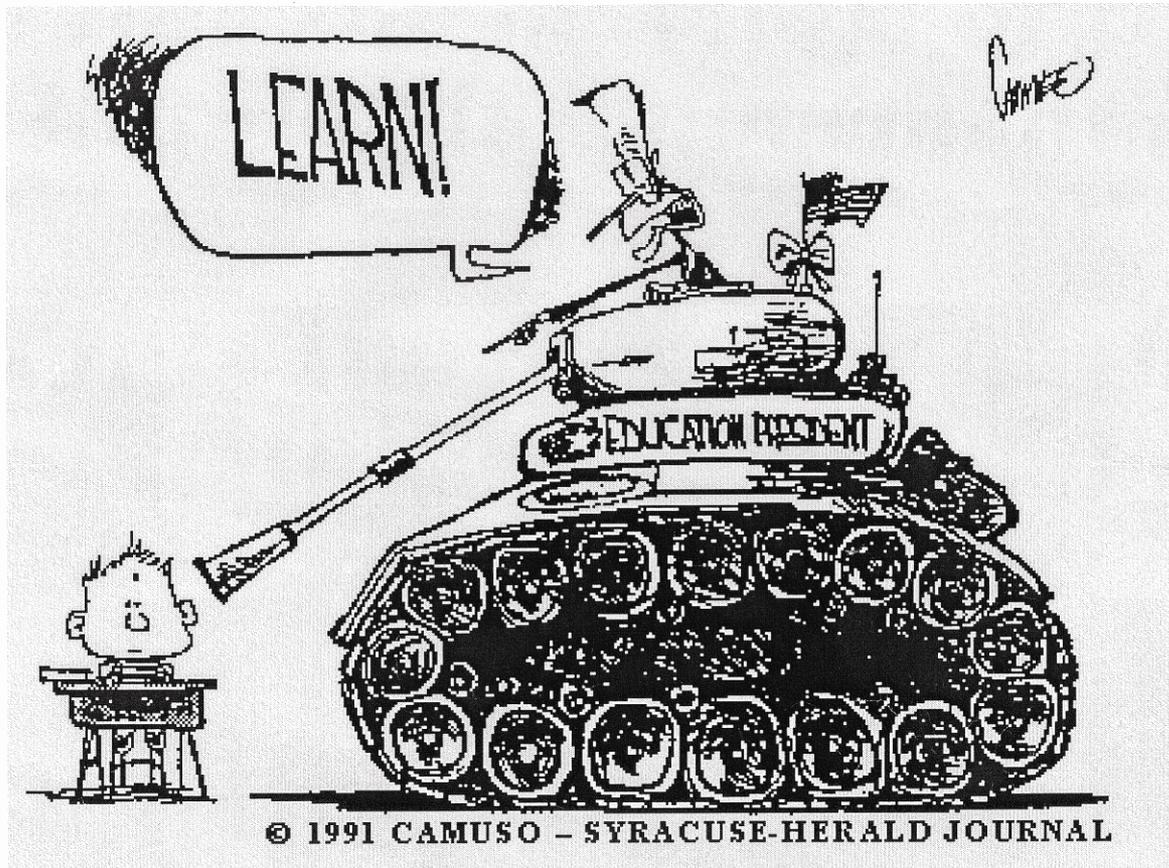
- System change is generally incremental and rarely radical
 - GCSE, Irish Junior Leaving Certificate, Hong Kong reform
- Competing policy agendas
 - 21st century learning vs accountability/selection

The 21st century learner agenda

1	Ways of thinking	<ul style="list-style-type: none">▶ Creativity & innovation▶ Critical thinking, problem-solving, decision-making▶ Learning to learn, metacognition
2	Ways of working	<ul style="list-style-type: none">▶ Communication▶ Collaboration
3	Tools for working	<ul style="list-style-type: none">▶ Information literacy▶ ICT literacy
4	Skills for living in the world	<ul style="list-style-type: none">▶ Local & global citizenship▶ Life & career▶ Personal & social responsibility

P. Griffin et al. (eds.), *Assessment and Teaching of 21st Century Skills*,
University of Melbourne, Australia

Is this the message of the “Standards” movement?



What kind of learners are we producing?

Attitudes to learning

- *Surface* – teacher dependent, ‘what do I need to do to pass?’
- *Strategic / instrumental* – the focus is on getting good marks/grades rather than on the learning (encouraged by low quality assessments)
- *Deep* – need to understand and make sense, thinking for themselves (the need for assessments that encourage the students to think)

What kind of learning do we want?

The case of Ruth

Learning the formula for each exam and practising it endlessly. I got an A1 in English because I knew exactly what was required in each question. I learned off the sample answers provided by the examiners and I knew how much information was required and in what format in every section of the paper. That's how you do well in these examinations...

There's no point in knowing about stuff that is not going to come up in the exams. I was always frustrated by teachers who would say 'You don't need to know this for the exams but I'll tell you anyway'. I wanted my A1 – what's the point of learning material that won't come up in the exams?

Is Ruth a good learner?

Brakes on change

The pull of:

- *Manageability*
Simpler, more cost-effective, risk to authenticity.
- *Standardisation* – common task, timing
Link to accountability and selection – less choice, less creativity?
- *Predictability*
The past paper tradition.

So what fit-for-the-future steps can we take?

David Boud's *double duty* of assessment:

Assessment activities...

have to focus on the immediate task and on implications for equipping students for lifelong learning in an unknown future

...they have to attend to both the process and the substantive domain.

So what fit-for-the-future steps can we take?

Five principles:

1. If teachers are going to teach to the test (and they will) then the test must encourage the skills and knowledge which the curriculum requires
2. The form of the test will influence teaching and learning – *a multiple-choice test of ‘knowledge in bits’ will lead to ‘teaching in bits’*
3. The predictability of a test will affect whether the teaching emphasis is on deep or surface learning approaches.
4. Tests must help to motivate those tested, through accessibility and fairness
5. How the results are interpreted and used are of critical importance.

(*Testing Times*, 2008)

“...the test must encourage the skills and knowledge which the curriculum requires”

Frederiksen and Collins *systemically valid test*:

One that induces in the education system curricular and instructional changes that foster the development of the cognitive skills that the test is designed to measure.

It is the *aims* of the course, rather than its content, that should determine the purpose and form of its assessment.

John White – National curriculum development in England

Ayres' study of 25 exceptional teachers

- Top 1% of exam results for over 6 years
- Their students do *worse* in other subjects

Findings:

1. Not exam driven – often chose difficult options and went beyond syllabus
2. Passion for their subject – teaching structured and teacher-led
3. Used whole-class question and answer
4. Emphasised holistic understanding & big picture
5. No spoon-feeding or answers during a task – students had to write their own notes.

(Ayres *et al* 2004)

2. *The form of the test will influence teaching and learning*

This is about fitness-for-purpose

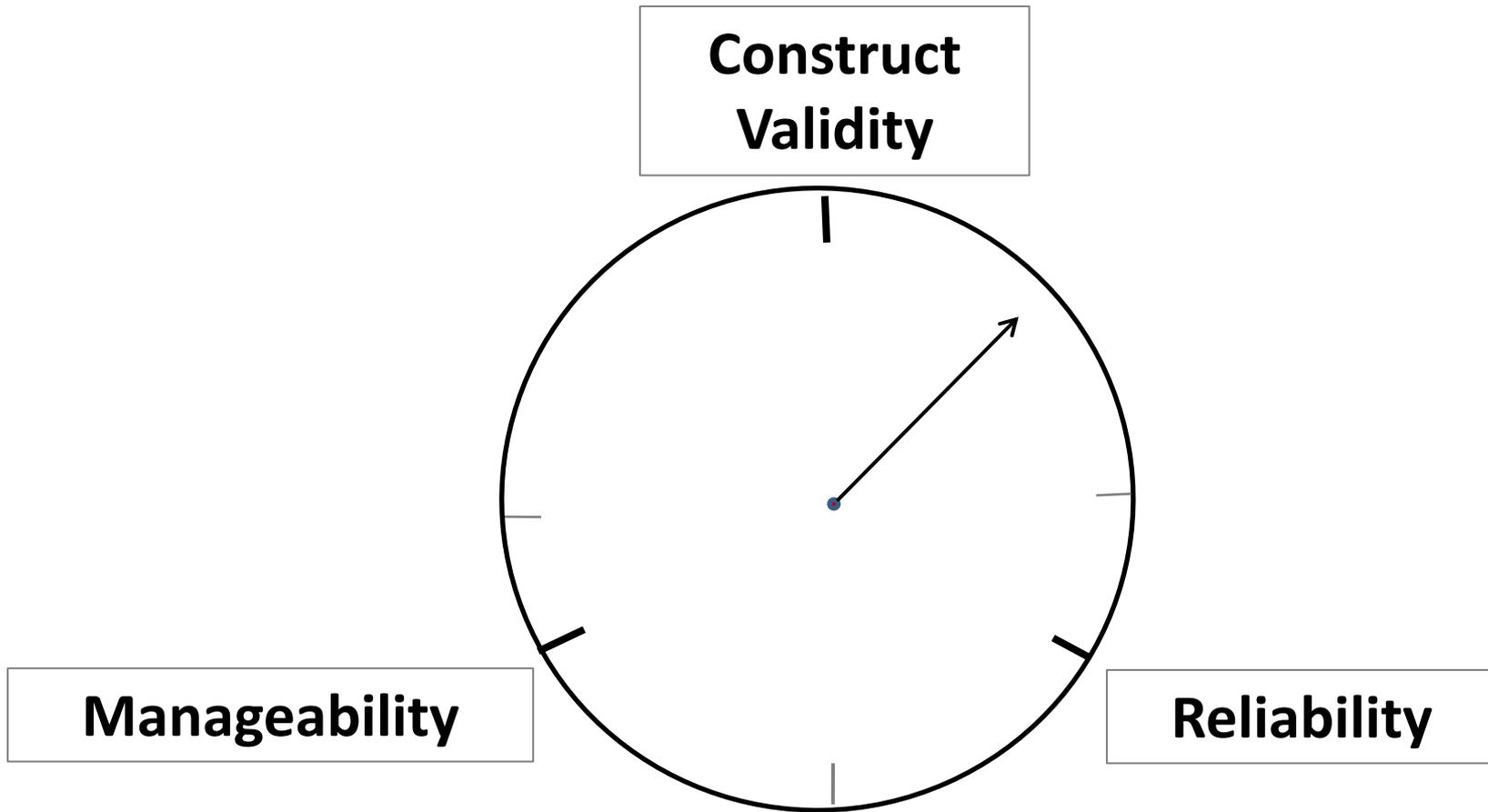
It's not just about format (m/c vs essay – Baumgart and Halse)

How the subject is *sampled* is a major threat to validity

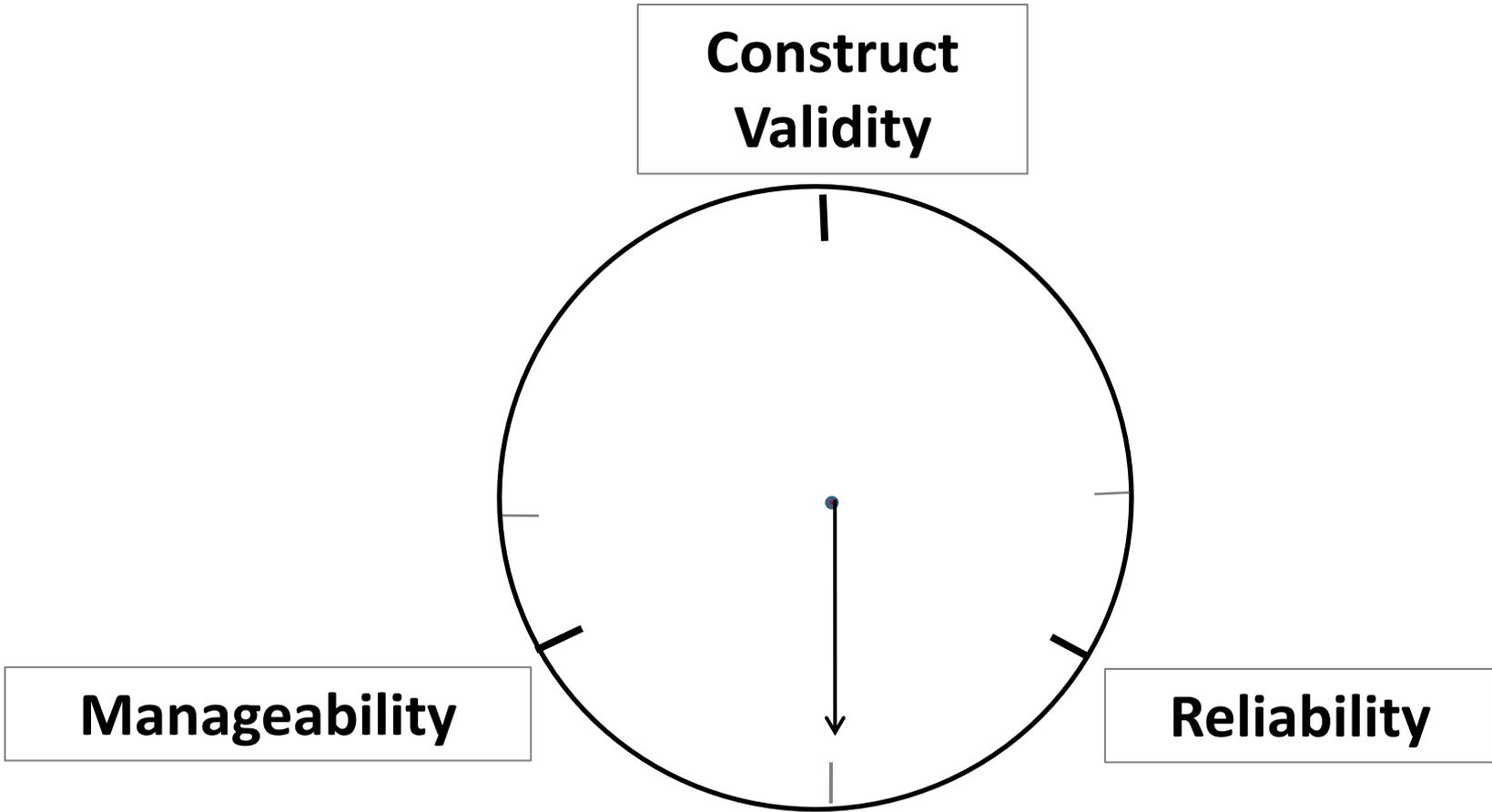
- English without Speaking & Listening (& extended writing at Key Stage 2 in England)
- Science without practicals?

Dependability – the optimal trade-off of validity, reliability and manageability

Dependability: The one-handed clock



Dependability: The one-handed clock



3...The predictability of a test will affect whether the teaching emphasis is on deep or surface learning approaches.

Some predictability is necessary – format etc.

The problems of the past paper tradition (and of specimen papers in new exams)

How do we reduce predictability?

- gradually & with notice (Eva Baker)
- predictable format – unpredictable topic
(Doug McCurry at the Australian Council for Educational Research (ACER))

The problem of predictability

Garrison Keillor offered this wry example in his *Lake Wobegon Days*:

For years, students of the senior class were required to read ['Phileopolis'] and answer questions about its meaning etc. Teachers were not required to do so, but simply marked according to the correct answers supplied by Miss Quoist, including: (1) To extend the benefits of civilization and religion to all peoples, (2) No, (3) Plato, and (4) A wilderness cannot satisfy the hunger for beauty and learning, once awakened.

The test was the same from year to year, and once the seniors found the answers and passed them to the juniors, nobody read 'Phileopolis' anymore.

ACT Test - 2013 *The Science of Happiness*

Read carefully the material on these two pages.

Produce a piece of writing of about 600 words, giving your point of view on the major issue raised in the material.

You need not refer to any of the material specifically, but you must deal with the major issue in it. Do not summarise the material.

Give your piece a clear title.

You should assume you are writing to persuade an intelligent adult audience.

Scripts of significantly more than 600 words will not be read in full.

CRITERIA FOR ASSESSMENT

- Quality of thought and content
- Structure and organisation
- Expression, style and mechanics.

WHAT THE EXAMINERS WILL LOOK FOR

- Your choice of an aspect of the major issue directly related to the material presented
- The coherence of your persuasive argument
- The quality of your evidence or supporting arguments
- The control of language and form, including clarity, precision and appropriateness for the target audience.

SOME THINGS YOU MIGHT CONSIDER BEFORE HANDING IN YOUR PIECE OF WRITING

- Have I given my work a clear and accurate title?
- Does the work have a clear organisation and structure?
- Have I read over the piece of writing and edited the final copy for spelling, punctuation and sentence structure?
- Are there places in which my handwriting would be difficult to follow and should be tidied?
- Are all the alterations I have made clear to the reader?

Question 8

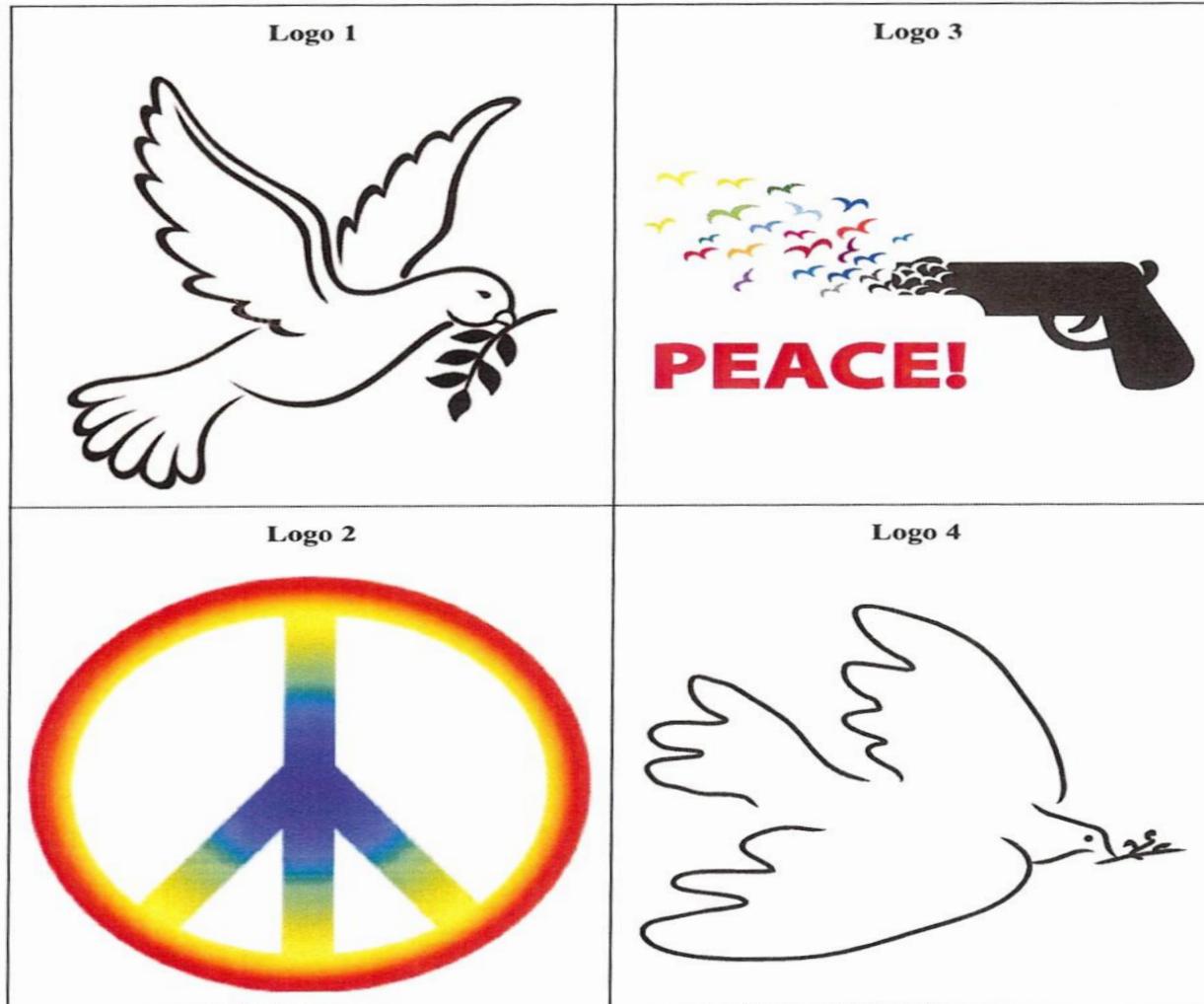
Compare the four logos symbolising peace below.

Which logo do you find most effective and why?

Your responses will be judged on the:

- accuracy of the description and understanding of the material; and
- substance and quality of the interpretation offered.

(3)



4. Tests must help to motivate those tested, through accessibility and fairness.

Access questions	Curricular questions	Assessment questions
Who gets taught and by whom?	Whose knowledge is taught?	What knowledge is assessed and equated with achievement?
Are there differences in the resources available for different groups?	Why is it taught in a particular way to this particular group?	Are the form, content and mode of assessment appropriate for different groups and individuals?
What is incorporated from the cultures of those attending?	How do we enable the histories and cultures of people of color, and of women, to be taught in responsible and responsive ways? Apple 1989	Is this range of cultural knowledge reflected in definitions of achievement? How does cultural knowledge mediate individuals' responses to assessment in ways which alter the construct being assessed? Gipps and Murphy 1994

So where next?

- Classroom assessment can lead the way
 - we ‘have to focus on the immediate task and on implications for equipping students for lifelong learning in an unknown future’
 - We have better opportunities to encourage deep and flexible learning
- New technology can show what is possible – and how it is manageable
- But we can’t sit back and wait for a technological miracle – we have to change the culture.

