E-Assessment for Learning: balancing challenges and opportunities

Denise Whitelock
The Open University, Walton Hall,
Milton Keynes MK7 6AA
d.m.whitelock@open.ac.uk
Assessment
I hate marking but want the tasks and feedback to assist student learning
The e-Assessment Challenge

• Constructivist Learning – Push

• Institutional reliability and accountability – Pull
MCQs: Variation on a theme

Example of LAPT Certainty-Based Marking, UK cabinet ministers demo exercise showing feedback, University College, Tony Gardner-Medwin

Drug Chart Errors and Omissions, Medicines Administration Assessment, Chesterfield Royal Hospital
Example of practice Thinking Skills Assessment" (TSA) question, Admissions Interview, Cambridge Assessment, Steve Lay

Example of practice Thinking Skills Assessment" (TSA) feedback, Admissions Interview, Cambridge Assessment, Steve Lay
Scaffolding and High Stakes assessment

- Math for Science
- Tutor less course
- Competency led
- No point to cheat
- Web home exam
- Invigilation technologies
Self diagnosis

• Basic IT skills, first year med students (Sieber, 2009)
• Competency based testing
• Repeating tests for revision
• Enables remedial intervention
Theoretical drivers for Assessment

- Piaget (1930) Individual manipulating surroundings -- mental representations can be tested Induction of rules and their testing ...Nuffield science
- Cognitive pyschology( 70s 80s) How are these mental representations stored? Classification development and how this can go wrong. .... Misconceptions and mal rules....... AI
- Bruner (1982) assessment tasks will match competency levels depending on level of help...SCAFFOLDING
Scaffolding and Interactivity; OU Science Foundation Course

- Interaction, Feedback loops
  - Tell, Explore, Check
  - Predict, Look and Explain
- Entering the discourse of a subject via audio feedback
- Scaffolded text feedback (Bruner & Woods)
- SHOW ME button
Interactive Tasks

- Games
- Simulations
- Making the abstract concrete
- Directing the sequence of an animation
Interactivity and Cognitive Change Scores

- Galapagos
- Organic Molecules
- Meiosis & Mitosis
- Seismic Waves

- Interactivity
- Cog. Change
Theoretical drivers: Social Constructivism

- Vygotsky (1978) individuals shape cultural settings which shapes minds no longer individual

- Activity theory (Engstrom 1987) Tool mediation

- Situated Cognition (Lave and Wenger 1991) Authentic assessment

- Peer interaction and assessment

- Learning conversations Laurillard (2002)
<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Descriptor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Authentic</td>
<td>Involving real-world knowledge and skills</td>
</tr>
<tr>
<td>Personalised</td>
<td>Tailored to the knowledge, skills and interests of each student</td>
</tr>
<tr>
<td>Negotiated</td>
<td>Agreed between the learner and the teacher</td>
</tr>
<tr>
<td>Engaging</td>
<td>Involving the personal interests of the students</td>
</tr>
<tr>
<td>Recognise existing skills</td>
<td>Willing to accredit the student’s existing work</td>
</tr>
<tr>
<td>Deep</td>
<td>Assessing deep knowledge – not memorization</td>
</tr>
<tr>
<td>Problem oriented</td>
<td>Original tasks requiring genuine problem solving skills</td>
</tr>
<tr>
<td>Collaboratively produced</td>
<td>Produced in partnership with fellow students</td>
</tr>
<tr>
<td>Peer and self assessed</td>
<td>Involving self reflection and peer review</td>
</tr>
<tr>
<td>Tool supported</td>
<td>Encouraging the use of ICT</td>
</tr>
</tbody>
</table>

Elliott’s characteristics of Assessment 2.0 activities
Authentic assessments - e-portfolios

Electronic NVQ portfolio cover contents page, OCR IT Practitioner, EAIHFE, Robert Wilsdon
Building e-portfolios on a chef’s course

food preparation for e-portfolio, Modern Apprenticeship in Hospitality and Catering, West Suffolk College, Mike Mulvihill

Evidence of food preparation skill for e-portfolio, Modern Apprenticeship in Hospitality and Catering, West Suffolk College, Mike Mulvihill
Sharing e-portfolios: The Netfolio concept

- Social constructivism
- Connecting e-portfolios (Barbera, 2009)
- Share and build upon a joint body of evidence
- Trialled with 31 PhD students at a virtual university
- Control group used but Netfolio group obtained higher grades
- Greater visibility of revision process and peer assessment in the Netfolio system
Peer Assessment and the WebPA Tool

- Loughborough (Loddington et al, 2009)
- Self assess and peer assess with given criteria
- Group mark awarded by tutor
- Students rated:
  - More timely feedback
  - Reflection
  - Fair rewards for hard work
- Staff rated:
  - Time savings
  - Administrative gains
  - Automatic calculation
  - Students have faith in the administrative system
Mobile Technologies and Assessment

- MCQs, PDAs
  Valdiva & Nussbaum (2009)
- Polls, instant surveys
- Simpson & Oliver (2007)
- Draper (2009) EVS
Gains from Formative Assessment

- Mean effect size on standardised tests between 0.4 to 0.7 (Black & Williams, 1998)
- Particularly effective for students who have not done well at school
  http://kn.open.ac.uk/document.cfm?docid=10817
- Can keep students to timescale and motivate them
- How can we support our students to become more reflective learners and engage in formative assessment tasks?
Formative assessment in Level 3 Physics and Astronomy

S382 Astrophysics 2010
S383 The Relativistic Universe 2010

Black = Summative

Red = Formative
Formative assessment strategy

SM358 and SMT359: Quantum and Electromagnetism

iCMA Maths

iCMA 51
iCMA 52
iCMA 53
iCMA 54
iCMA 55
iCMA 56

3 Revision Medleys

TMA 01
TMA 02
TMA 03
TMA 04

OCAS Threshold

30% or more on 7 or more assignments including at least 2 TMAs.

Course Grade determined by exam score alone.
# Formative assessment strategy

**S382: Astrophysics**

<table>
<thead>
<tr>
<th>iCMA 51</th>
<th>TMA 01</th>
</tr>
</thead>
<tbody>
<tr>
<td>iCMA 52</td>
<td>TMA 02</td>
</tr>
<tr>
<td>iCMA 53</td>
<td>TMA 03</td>
</tr>
<tr>
<td>iCMA 54</td>
<td>Project</td>
</tr>
<tr>
<td>iCMA 55</td>
<td>TMA 04</td>
</tr>
<tr>
<td></td>
<td>TMA 05</td>
</tr>
</tbody>
</table>

**OCAS Threshold**

30% or more on 8 or more assignments.

TMAs and iCMAs cover complementary material.

**Course Grade determined by exam (2/3)+ Project (1/3)**
Formative assessment strategy

S383 Relativistic Universe

iCMA 51   TMA 01
iCMA 52   TMA 02
iCMA 53   TMA 03
iCMA 54   EC1
iCMA 55   TMA 04
iCMA 56   TMA 05
TMA 06

OCAS Threshold

30% or more on 9 or more assignments.

TMAs and iCMAs cover complementary material.

Course Grade determined by exam \( (2/3) + \text{EC1} \ (1/3) \).
Attempts at TMAs (Physics)

SM358 2009
SM358 2010
SMT359 2008
SMT359 2010
Average marks for TMAs (Physics)

Blue = summative
Red = formative
Average marks for TMAs (Astronomy)

Blue = summative
Red = formative

S381/S382

S357/S383
iCMA marks (SM358)

SM358 iCMA scores 2009 and 2010

Red = 2010
Blue = 2009
Time of first attempts at iCMAs

e.g.

Cumulative total

% 100

iCMA52 (164)

Red = 2010
Blue = 2009

Recommended completion date

Exam 2010
Exam 2009
## The bottom line

<table>
<thead>
<tr>
<th></th>
<th>SM358 2009</th>
<th>SM358 2010</th>
<th>SMT359 2009</th>
<th>SMT359 2010</th>
</tr>
</thead>
<tbody>
<tr>
<td>Retention</td>
<td>70.7%</td>
<td>70.5%</td>
<td>62.7%</td>
<td>60.8%</td>
</tr>
<tr>
<td>Grade 1</td>
<td>22.0 %</td>
<td>22.7 %</td>
<td>12.9 %</td>
<td>7.3 %</td>
</tr>
<tr>
<td>Grade 2</td>
<td>15.2 %</td>
<td>13.5 %</td>
<td>9.8 %</td>
<td>10.5 %</td>
</tr>
<tr>
<td>Grade 3</td>
<td>12.3 %</td>
<td>17.9 %</td>
<td>15.1 %</td>
<td>17.1 %</td>
</tr>
<tr>
<td>Grade 4</td>
<td>11.5 %</td>
<td>6.3 %</td>
<td>12.3 %</td>
<td>13.3 %</td>
</tr>
<tr>
<td>Grade Q/5</td>
<td>9.7 %</td>
<td>10.1 %</td>
<td>12.6%</td>
<td>12.6%</td>
</tr>
<tr>
<td>Pass rate</td>
<td>61.0 %</td>
<td>60.4 %</td>
<td>50.1%</td>
<td>48.2%</td>
</tr>
</tbody>
</table>
## The bottom line

<table>
<thead>
<tr>
<th></th>
<th>S381 2008</th>
<th>S382 2010</th>
<th>S357 2009</th>
<th>S383 2010</th>
</tr>
</thead>
<tbody>
<tr>
<td>Retention</td>
<td>59.2%</td>
<td>57.5%</td>
<td>60.1%</td>
<td>56.4%</td>
</tr>
<tr>
<td>Grade 1</td>
<td>10.1%</td>
<td>12.1%</td>
<td>12.2%</td>
<td>9.9%</td>
</tr>
<tr>
<td>Grade 2</td>
<td>24.7%</td>
<td>18.9%</td>
<td>6.1%</td>
<td>13.7%</td>
</tr>
<tr>
<td>Grade 3</td>
<td>13.5%</td>
<td>15.2%</td>
<td>17.0%</td>
<td>20.6%</td>
</tr>
<tr>
<td>Grade 4</td>
<td>4.7%</td>
<td>6.8%</td>
<td>17.4%</td>
<td>6.1%</td>
</tr>
<tr>
<td>Grade Q/5</td>
<td>6.2%</td>
<td>4.5%</td>
<td>7.4%</td>
<td>6.1%</td>
</tr>
<tr>
<td>Pass rate</td>
<td>53.0%</td>
<td>53.0%</td>
<td>52.7%</td>
<td>50.3%</td>
</tr>
</tbody>
</table>
Conclusions

• Formative assessment makes much less difference than we might have imagined.

• In the first year of running this strategy, the results are comparable with (or marginally worse than) those for summative assessment.

• BUT we are still learning how to optimise the strategy, and there is the potential for future improvements.
Physicists' conclusions

- Formative assessment makes much less difference imagined.
- In the first year of running this strategy, the results are comparable with (or marginally worse than) those for summative assessment.
- BUT we are still learning how to optimise the strategy, and there is the potential for future improvements.
MU 123: Discovering Mathematics Tim Lowe

- 30 point introductory module
- Launched February 2010
- 2 presentations a year
- 1800 students oer presentation
## 2 Types of Computer Assisted Assessment quizzes

<table>
<thead>
<tr>
<th>Practice Quizzes</th>
<th>iCMAs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Formative</td>
<td>Summative</td>
</tr>
<tr>
<td>Immediate feedback</td>
<td>Delayed feedback</td>
</tr>
<tr>
<td>3 tries per question</td>
<td>Single attempt</td>
</tr>
<tr>
<td>Worked solutions</td>
<td>6 questions per unit</td>
</tr>
<tr>
<td>Can repeat</td>
<td>12% of grade</td>
</tr>
<tr>
<td>12 questions per unit</td>
<td></td>
</tr>
</tbody>
</table>
Grades

Average best finalised grade by icma/pq

Grade (%)
Practice Quiz repeats

Numbers of students taking multiple attempts at a PQ

Number of attempts at a PQ

Number of students
End of course survey 2010

- The computer-based individual learning activities for this course supported my learning
- 92.2% agree 0.5% disagree
- Increased retention
Collaborative formative assessment with Global Warming

Balancing sensitivities

- **Solar constant**: 1368.00 Wm⁻²
- **CO₂**: 350.00 ppm
- **H₂O**: 30.55 kgm⁻²
- **Aerosol content**: 0.00 Pin
- **Albedo**: 0.087
- **Ice and snow**: 0.100
- **Cloud cover**: 0.505

Change in GMST/°C

All factors are set again to their early 1990s values. You can now change more than one factor at once. Try a few combinations of changes. (To reset to the early 1990s values press the Reset button.)

Click on the forward arrow to proceed.
Global Warming

Exploring sensitivities

- solar constant: 1368 W m$^{-2}$
- CO$_2$: 350 p.p.m.
- H$_2$O: 34.17 kg m$^{-2}$
- aerosol content: 0.00 Pin
- albedo: 0.087
- ice and snow: 0.100
- cloud cover: 0.504

To change a factor, move the mouse cursor over its slider knob and press the left mouse button. While holding this button down, move the mouse left or right. The changing value is displayed just to the right of the slider. Release the mouse button when you have the desired value. Try increasing the solar constant and press the Run button. Watch the changing GMST value and graph.
Next: ‘Yoked’ apps via BuddySpace

('yoked', but without full screen sharing required!)
Global Warming: Simlink Presentation

To change a factor, move the mouse cursor over its slider knob and press the left mouse button. While holding this button down, move the mouse left or right. The changing value is displayed just to the right of the slider. Release the mouse button when you have the desired value. Try increasing the solar constant and press the Run button. Watch the changing GMST value and graph.

Hi Simon, why don't you see what happens to GMST when you increase the amount of CO₂ in the atmosphere.
Research issues from CAA: Marking Free text

- LSA (Laudaurer, Laham & Foltz 2003)
- Christie SEAR (1999)
- Knowledge engineering Pulman & Sukkarieh (2005)
- IAT (Jordan and Mitchell 2009)
- Genre Analysis Moreale & Vargas – Vera (2005)
Open Comment addresses the problem of free text entry

- Automated formative assessment tool
- Free text entry for students
- Automated feedback and guidance
- Open questions, divergent assessment
- No marks awarded
- For use by Arts Faculty
Three Causal Question Types proved to be suitable for Open Comment

1. Analysis of statistics, usually presented to the student as a table
2. Comprehension of a set text
3. Identifying similarities and differences for a given event
Open Comment components

- A Java-based feedback system
- A web service shell
- A graphical interface for testing
- A Moodle-based question type
- A forms-based editing tool
Open Comment Arts

Moodle > AA000 > Quizzes > Combined Quiz > Attempt 1

Preview Combined Quiz

Read the first paragraph in which Joll analyses why the outbreak of the First World War has been studied in such detail. What reasons does he give?

Answer: No idea

Save without submitting  Submit all and finish
Open Comment Arts

Moodle > AA000 > Quizzes > Combined Quiz > Review

You are logged in as Stuart Watt (Logout)

Start again

Started on: Thursday, 3 July 2008, 01:16 pm
Completed on: Thursday, 3 July 2008, 01:17 pm
Time taken: 45 secs

Click here to go back to the course

1

Read the first paragraph in which Joll analyses why the outbreak of the First World War has been studied in such detail. What reasons does he give?

Answer: No idea

Feedback:

Maybe you are a bit confused by the question. It may be helpful to remember you are not being asked about the causes directly, but why the causes have been so extensively studied.
Preview Combined Quiz

1. Read the first paragraph in which Joll analyses why the outbreak of the First World War has been studied in such detail. What reasons does he give?

Answer:
The First World War was a turning point in History, showing that lessons might not be learnt. Recognising the causes of the war also to some extent vindicated the huge loss of life.
Open Comment Arts

Moodle > AA000 > Quizzes > Combined Quiz > Review

Click here to go back to the course

1. Read the first paragraph in which Joll analyses why the outbreak of the First World War has been studied in such detail. What reasons does he give?

Answer: The First World War was a turning point in History, showing that lessons might not be learnt. Recognising the causes of the war also to some extent vindicated the huge loss of life.

Feedback:

Well done on this submission

- This question does not ask about the causes directly, but why the causes have been so extensively studied. The issue of blame ought to form part of an answer to this question.
- This is absolutely right - what Joll is saying in technical terms, is that 1914 marked a turning point.
- You have done well here to use the term 'turning point', which is a technically accurate term within the discourse of the discipline.
Stages of analysis by computer of students’ free text entry for Open Comment: advice with respect to content (socio-emotional support stylised example)

- **STAGE 1a: DETECT ERRORS** E.g. Incorrect dates, facts. (Incorrect inferences and causality is dealt with below)
- Instead of concentrating on X, think about Y in order to answer this question. Recognise effort (Dweck) and encourage to have another go.
- You have done well to start answering this question but perhaps you misunderstood it. Instead of thinking about X which did not……… Consider Y.
Computer analysis continued

- **STAGE 2a: REVEAL FIRST OMISSION**
  - Consider the role of Z in your answer Praise what is correct and point out what is missing Good but now consider the role X plays in your answer

- **STAGE 2b: REVEAL SECOND OMISSION**
  - Consider the role of P in your answer Praise what is correct and point out what is missing Yes but also consider P. Would it have produced the same result if P is neglected?
Final stages of analysis

- STAGE 3: REQUEST CLARIFICATION OF KEY POINT 1
- STAGE 4: REQUEST FURTHER ANALYSIS OF KEY POINT 1 (Stages 3 and 4 repeated with all the key points)
- STAGE 5: REQUEST THE INFERENCE FROM THE ANALYSIS OF KEY POINT 1 IF IT IS MISSING
- STAGE 6: REQUEST THE INFERENCE FROM THE ANALYSIS OF KEY POINT 1 IF IT IS NOT COMPLETE
- STAGE 7: CHECK THE CAUSALITY
- STAGE 8: REQUEST ALL THE CAUSAL FACTORS ARE WEIGHTED
Where are we now?

• Opening up with Open Source
• Moving towards vision and not losing sight of it through tool adaptation
• More work to do for Arts
• Open Comment - pedagogical model open to test
• Feedback
  • Changing pedagogy
  • Another handle on misconceptions
Open Comment drivers for reflection

- Students are able to find facts similar to X
- Know how X might be disputed
- Are able to make predictions about X
- Know how to use X in an argument
- Know how far X can be pushed
- Supported with tools and strategies
LISC: Aily Fowler

- Kent University ab-initio Spanish module
  - Large student numbers
  - Skills-based course
  - Provision of sufficient formative assessment meant unmanageable marking loads
  - Impossible to provide immediate feedback
    - leading to fossilisation of errors
The LISC solution: developed by Ali Fowler

- A CALL system designed to enable students to:
  - Independently practise sentence translation
  - Receive immediate (and robust) feedback on all errors
  - Attend immediately to the feedback (before fossilisation can occur)
How is the final mark arrived at in the LISC System?

• The two submissions are *unequally* weighted
  • Best to give more weight to the first attempt
    • since this ensures that students give *careful* consideration to the construction of their first answer
    • but can improve their mark by refining the answer
  • The marks ratio can vary (depending on assessment/feedback type)
    • the more information given in the feedback, the lower the weight the second mark should carry
Heuristics for the final mark

- If the ratio is skewed too far in favour of the first attempt…
  - students are less inclined to try hard to correct non-perfect answers

If the ratio is skewed too far in favour of the second attempt…

- students exhibit less care over the construction of their initial answer
Feedback: Advice for Action

• Students must decode feedback and then act on it
  Boud (2000)

• Students must have the opportunity to act on feedback
  Sadler (1989)

• Gauging efficacy through student action
Welcome to Open Mentor

You're here: Open Mentor >> View reports >> Comments analysis

OpenMentor comment analysis - for individual courses

This page shows details of how each course assignment's comments have been analysed. If you would like to see further detail of the analysis, please click on the bar, pie or table chart links.

The important features to be looking for in the graph are the size of the blue bars above and below the zero line. If there is a block below the line for a given category, for example category A, this means that not enough praise was given to the student on the assignment. If there is a block above the line then too many comments have been given in this category for the mark awarded. For example, if there is a large block above the line in category D, then too many negative comments have been made on the assignment.

OpenMentor clustered chart

Categories

- [Ideal] - [Actual]
<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Descriptor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Authentic</td>
<td>Involving real-world knowledge and skills</td>
</tr>
<tr>
<td>Personalised</td>
<td>Tailored to the knowledge, skills and interests of each student</td>
</tr>
<tr>
<td>Negotiated</td>
<td>Agreed between the learner and the teacher</td>
</tr>
<tr>
<td>Engaging</td>
<td>Involving the personal interests of the students</td>
</tr>
<tr>
<td>Recognise existing skills</td>
<td>Willing to accredit the student’s existing work</td>
</tr>
<tr>
<td>Deep</td>
<td>Assessing deep knowledge – not memorization</td>
</tr>
<tr>
<td>Problem oriented</td>
<td>Original tasks requiring genuine problem solving skills</td>
</tr>
<tr>
<td>Collaboratively produced</td>
<td>Produced in partnership with fellow students</td>
</tr>
<tr>
<td>Peer and self assessed</td>
<td>Involving self reflection and peer review</td>
</tr>
<tr>
<td>Tool supported</td>
<td>Encouraging the use of ICT</td>
</tr>
</tbody>
</table>

**Advice for Action**

Elliott’s characteristics of Assessment 2.0 activities
Embedding Assessment into Pedagogical Frameworks

• Luca & Oliver (2002)

• Boud & Falchikov (2007)

• Bartlett-Bragg (2008)
Cycles in e-Assessment

Cycles of e-assessment

Motivation

Design

Creation

Testing

Delivery

Data retrieval, processing & feedback

Evaluation of outcomes

Feedback into revision of process, content and wider context
The 4Ts Pyramid

- Tool Development
- Training of staff
- Transformation Tasks
- Transfer Learning
e-Assessment: effectiveness vs. innovation
Scatter diagram to illustrate relationship between Maturity vs Availability

- OU
- Surrey
- Heriot Watt
- Dundee
- Plymouth
- Southampton
- Loughborough
- Ulster
- UCL
- Glamorgan
- Derby
- Birkbeck
- Plymouth
- Cambridgeshire Assessment
- Warwickshire College
- Coleg Sir Gar
- East Antrium
- Belfast
- West Suffolk
- Calibrand

Maturity vs Availability

14 16 18 20 22 24

12 14 16 18 20 22 24
Training

- Question development
- Quality REAQ
- Plagiarism software
- Advice for Action
- Socio-emotive content
- Maintain empathy with the Learner
e-Assessment Futures

- Research issues
- Adaptive testing
- Automatic marking
- Learning Analytics
- Data mining
- Web 2.0, 3.0 ...
Rearranging

- Promoting champions
- Multidisciplinary teams
- Open source
- Pedagogical frameworks
National Union of Students’ Principles of Effective Assessment  
Times Higher Education, 29th January 2009

- Should be for learning, not simply of learning
- Should be reliable, valid, fair and consistent
- Should consist of effective and constructive feedback
- Should be innovative and have the capacity to inspire and motivate.
- Should be conducted throughout the course, rather than being positioned as a final event
- Should develop key skills such as peer and reflective assessment
References


Three Assessment Special Issues

