The black box of tertiary assessment

John Hattie
Visible Learning Laboratories
University of Auckland

Symposium on Tertiary Assessment and Higher Education Student Outcomes: Policy, Practice, and Research
A revolution in assessment …

- Assessment for Learning
- NCEA & its standards based approach
- Emphasis on reporting more than scoring
- Peer collaborative assessment
- Learning intentions and success criteria
- Realization of the power of feedback
- Constructive alignment of learning & outcomes
Inside the black box
1. The multiple outcomes

I. Achieving competence

II. Managing emotions – from those that interfere with learning (anger, anxiety, hopelessness – to those that assist (optimism, hopefulness).

III. Mature interpersonal relations – respecting differences, working with peers

IV. Moving from autonomy to independence – moving from needing assurance and approval of others to self-sufficiency, problem solving, and making decisions

V. Establishing identity – self-esteem and self-efficacy

VI. Developing purpose – from Who am I? and Where am I? to Where am I going?

VII. Developing integrity
1. **Generic skills**

- critical thinking
- analytic reasoning
- problem-solving
- written communication skills
- generation of knowledge
- interaction between substantive and methodological expertise
Assessing Higher Education Learning Outcomes (AHELO) - OECD

1. Generic skills
2. Discipline-specific skills

engineering and economics.
1. Generic skills
2. Discipline-specific skills
3. Student outcomes

- absolute performance or raw scores of students
- a measure of incremental learning (or “value-added”)
Assessing Higher Education Learning Outcomes (AHELO) - OECD

1. Generic skills
2. Discipline-specific skills
3. Student outcomes
4. Contextual measures
   - Academic studies and teaching (contact between students, counseling, courses offered, opportunities for e-learning, study organization and teaching evaluation);
   - Equipment
   - International orientation
   - Job market and career orientation
   - Research
   - Study location and TEI
   - Overall opinions
<table>
<thead>
<tr>
<th>Rank</th>
<th>Research</th>
<th>Research + Teaching</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>‘08</td>
<td>‘07</td>
</tr>
<tr>
<td>1</td>
<td>Harvard</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>Stanford</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>Univ California – Berkeley</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>Cambridge</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>MIT</td>
<td>5</td>
</tr>
<tr>
<td>6</td>
<td>California Inst Tech</td>
<td>6</td>
</tr>
<tr>
<td>7</td>
<td>Columbia</td>
<td>7</td>
</tr>
<tr>
<td>8</td>
<td>Princeton</td>
<td>8</td>
</tr>
<tr>
<td>9</td>
<td>Univ of Chicago</td>
<td>9</td>
</tr>
<tr>
<td>10</td>
<td>Oxford</td>
<td>10</td>
</tr>
</tbody>
</table>
2 Constructive Alignment

- Learning and teaching activities
  - Designed to meet learning outcomes

- Intended Learning Outcomes

- Assessment methods
  - Designed to assess learning outcomes
3. What works best?

- 800 + meta-analyses
- 50,000 studies
- 240m students
Influences on Achievement?

- Decreased
- Zero
- Enhanced
Distribution of effects
Major conclusions

Almost everything works

Setting the bar at zero is absurd

Set the bar at $d = .40$

What works in schools, also works in Universities
<table>
<thead>
<tr>
<th>Rank</th>
<th>Influence</th>
<th>ES</th>
</tr>
</thead>
<tbody>
<tr>
<td>70</td>
<td>Time on Task</td>
<td>.38</td>
</tr>
<tr>
<td>71</td>
<td>Computer assisted instruction</td>
<td>.37</td>
</tr>
<tr>
<td>79</td>
<td>Frequent/Effects of testing</td>
<td>.34</td>
</tr>
<tr>
<td>103</td>
<td>Teaching test taking</td>
<td>.22</td>
</tr>
<tr>
<td>104</td>
<td>Visual/Audio-visual methods</td>
<td>.22</td>
</tr>
<tr>
<td>106</td>
<td>Class size</td>
<td>.21</td>
</tr>
<tr>
<td>111</td>
<td>Co-/Team teaching</td>
<td>.19</td>
</tr>
<tr>
<td>112</td>
<td>Web based learning</td>
<td>.18</td>
</tr>
<tr>
<td>120</td>
<td>Mentoring</td>
<td>.15</td>
</tr>
<tr>
<td>122</td>
<td>Gender</td>
<td>.12</td>
</tr>
<tr>
<td>126</td>
<td>Distance Education</td>
<td>.09</td>
</tr>
<tr>
<td>130</td>
<td>College halls of residence</td>
<td>.05</td>
</tr>
<tr>
<td>Rank</td>
<td>Influence</td>
<td>ES</td>
</tr>
<tr>
<td>------</td>
<td>---------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>24</td>
<td>Cooperative vs. individualistic learning</td>
<td>.59</td>
</tr>
<tr>
<td>25</td>
<td>Study skills</td>
<td>.59</td>
</tr>
<tr>
<td>29</td>
<td>Mastery learning</td>
<td>.58</td>
</tr>
<tr>
<td>30</td>
<td>Worked examples</td>
<td>.57</td>
</tr>
<tr>
<td>34</td>
<td>Goals - difficulty</td>
<td>.56</td>
</tr>
<tr>
<td>36</td>
<td>Peer tutoring</td>
<td>.55</td>
</tr>
<tr>
<td>37</td>
<td>Cooperative vs. competitive learning</td>
<td>.54</td>
</tr>
<tr>
<td>48</td>
<td>Small group learning</td>
<td>.49</td>
</tr>
<tr>
<td>49</td>
<td>Concentration/Persistence/ Engagement</td>
<td>.48</td>
</tr>
<tr>
<td>56</td>
<td>Quality of Teaching</td>
<td>.44</td>
</tr>
<tr>
<td>63</td>
<td>Cooperative learning</td>
<td>.41</td>
</tr>
<tr>
<td>Rank</td>
<td>Influence</td>
<td>ES</td>
</tr>
<tr>
<td>------</td>
<td>--------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>1</td>
<td>Self-report grades</td>
<td>1.44</td>
</tr>
<tr>
<td>3</td>
<td>Providing formative evaluation to lecturers</td>
<td>.90</td>
</tr>
<tr>
<td>8</td>
<td>Teacher clarity</td>
<td>.75</td>
</tr>
<tr>
<td>9</td>
<td>Reciprocal teaching</td>
<td>.74</td>
</tr>
<tr>
<td>10</td>
<td>Feedback</td>
<td>.73</td>
</tr>
<tr>
<td>12</td>
<td>Spaced vs. Mass Practice</td>
<td>.71</td>
</tr>
<tr>
<td>13</td>
<td>Meta-cognitive strategies</td>
<td>.69</td>
</tr>
<tr>
<td>17</td>
<td>Creativity Programs</td>
<td>.65</td>
</tr>
<tr>
<td>18</td>
<td>Self-verbalization/Self-questioning</td>
<td>.64</td>
</tr>
<tr>
<td>19</td>
<td>Professional development</td>
<td>.62</td>
</tr>
<tr>
<td>20</td>
<td>Problem solving teaching</td>
<td>.61</td>
</tr>
</tbody>
</table>
Visible teaching & Visible learning

- What some lecturers do!
  - In active, calculated and meaningful ways
  - Providing multiple opportunities & alternatives
  - Teaching learning strategies
  - Around surface and deep learning
  - That leads to students constructing learning
Visible Teaching – Visible Learning

When teachers SEE learning through the eyes of the student

and

When students SEE themselves as their own teachers
4. Assessment for learning/
   Feedback from assessment
Feedback is information provided by an agent (e.g., teacher, peer, book, parent, self/experience) regarding aspects of one’s performance or understanding.
Feedback is evidence about:

- Where am I going?
- How am I going?
- Where to next?
The power of Feedback

![Diagram showing the power of feedback with different categories including Teacher Effects, Developmental Effects, Reverse Effects, and a zone of desired effects. The feedback value is denoted as d = .73.]

### KEY

- **Standard error**: .061 (Medium)
- **Rank**: 10th
- **Number of Meta-analyses**: 23
- **Number of Studies**: 1,287
- **Number of Effects**: 2,050
- **Number of People (10)**: 67,931
Is it ....?

- feedback as something *teachers provide to students*
Is it ....?

- feedback as something *teachers provide to students*

**NO NO NO NO NO** –

**IT IS ...**

- feedback is most powerful when it is from the *student to the teacher*
Feedback to teachers helps make learning visible

When teachers seek,
or at least are open to,
feedback from students as to what students know,
what they understand,
where they make errors,
when they have misconceptions,
when they are not engaged

—then teaching and learning can be synchronized and powerful
The key to feedback

- is when feedback that is received and acted upon by students

- many teachers claim they provide ample amounts of feedback but the issue is whether students receive and interpret the information in the feedback (Carless, 2006)

- At best, each student receives moments of feedback in a single day -- and not much from too many assignments

- Most feedback comes from peers, .... and ...
Feedback from assessment

- The role of scoring rubrics
  - Learning intentions and success criteria
- The beginning of computer based essay scoring
- The use of peer critique
- The power of peer assessment
- The use of peer collaboration
- Assessment for learning as well as of, and as learning
- Multiple opportunities + spaced practice
## 5 Assessment to get into University

### Prior meta-analyses

<table>
<thead>
<tr>
<th>Author</th>
<th>Year</th>
<th>Studies</th>
<th>r</th>
</tr>
</thead>
<tbody>
<tr>
<td>Goldberg &amp; Alliger</td>
<td>1992</td>
<td>10</td>
<td>.15</td>
</tr>
<tr>
<td>Morrison &amp; Morrison</td>
<td>1995</td>
<td>22</td>
<td>.22-.28</td>
</tr>
<tr>
<td>Kuncel, Hezlett, &amp; Ones</td>
<td>2001</td>
<td>1753</td>
<td>.13-.38</td>
</tr>
<tr>
<td><strong>Overall</strong></td>
<td></td>
<td></td>
<td><strong>.20-.35</strong></td>
</tr>
</tbody>
</table>
The two systems

NCEA
-- no. credits (quantity)
-- GPA (E=4, M=3, A=2, NA=0)

University approved only

Cambridge
-- cumulative weighted score
Cambridge

- CIE and GPA  \( r = .30 \)
Correlations with 1st year GPA …

- Cambridge  
  .30

- Credit-based NCEA model with University GPA  
  .52

- GPA NCEA and University GPA  
  .66

Thus, NCEA is 4.8 times \((.66^2/.30^2)\) more effective than CIE
Across Degrees

Credit
Let’s re-work the black box …

- NCEA mimics 1st year = ongoing assessments involving
  - a variety of tasks throughout the year
  - an increasing higher level of independence in producing projects or assessment tasks,
  - together with a final examination

- Cambridge typical of summative high school tests

- Bring on assessment for learning
  - Feedback from assessment
The black box of tertiary assessment

Thank you ...

j.hattie@auckland.ac.nz

www.education.auckland.ac.nz/staff/j.hattie/

www.visiblelearning.co.nz
VISIBLE LEARNING: A SYNTHESIS OF OVER 800 META-ANALYSES IN EDUCATION

The black box of tertiary assessment

j.hattie@auckland.ac.nz

www.education.auckland.ac.nz/staff/j.hattie/

www.visiblelearning.co.nz