Just how good are student-generated assessment questions?

Simon Bates
CWSEI Seminar Oct 16th 2012
Overview

I. About PeerWise

II. Pilot use at UoE 2010—

III. Question quality?

IV. Community, future research
I. About PeerWise
Paul Denny
Department of Computer Science
The University of Auckland
New Zealand
• Web-based Multiple Choice Question repository built by students

• Students:
  – develop new questions with associated explanations
  – answer existing questions and rate them for quality and difficulty
  – take part in discussions
  – can follow other authors

peerwise.cs.auckland.ac.nz
As a question author.....
Alternatives

Write up to five alternative answers for the question you have written above. Make sure each alternative is distinct, and of course, you must ensure that exactly one of the alternatives is the correct answer to your question. You may choose to define fewer than five alternatives (by simply leaving some of the text areas empty), but you must at least provide two alternatives.

You must indicate which of the alternatives is the correct answer to your question by selecting the letter to the left of the alternative.
**Explanation**

You should provide an explanation for your answer. This explanation will only be shown to people after they have selected what they think is the answer to your question, and may help to explain to them why the alternative you have suggested is indeed the correct answer.

**Topics**

You may define up to FIVE topics which are relevant to this question. These topic definitions will make it easier for everyone to find questions on certain topics.

**Existing topics:** You can select from the current list of topics:

- Acceleration
- Angular Velocity
- Balancing forces
- Basic Normal Force
- Beer
- Bugatti Veyron
- Centripetal Force
- Electrostatic forces
- Energy
- Fictitious Forces
- Forces and Motion
- Friction
- Galilean transforms
- Gravitational Force
- Kinetic Energy
- Momentum
- Newton
- Numbers
- Pendulum
- Periods of orbit
- Potential Energy
- SHM
- Science fiction
- Sound
- Space and Time
- Springs
- Tension
- Units
As a question answerer.....

**Physics 1A 2010-11**

**Question stats**

This question has been answered by 45 people and has an average rating of 4.00 (based on 33 ratings).

The answer suggested by the author of this question is the most popular answer.

**Answer the following question**

A guy was standing on a high cliff in a very very cold winter.
(In that case, the speed of sound is nearly 320 ms⁻¹)

He dropped 5 bells from the cliff to the ground in order to measure the height of the cliff.
(drop only one bell each time, and initial speed of the bell is 0 ms⁻¹)
<table>
<thead>
<tr>
<th>Time taken</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>First time</td>
<td>8.8 s</td>
</tr>
<tr>
<td>Second time</td>
<td>9.5 s</td>
</tr>
<tr>
<td>Third time</td>
<td>8.7 s</td>
</tr>
<tr>
<td>Fourth time</td>
<td>9.2 s</td>
</tr>
<tr>
<td>Fifth time</td>
<td>8.8 s</td>
</tr>
</tbody>
</table>

Calculate the height of the cliff. (Take $g=10 \text{ m/s}^2$)

**Select your answer:**

<table>
<thead>
<tr>
<th>OPTION</th>
<th>ALTERNATIVE</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>100 m</td>
</tr>
<tr>
<td>B</td>
<td>300 m</td>
</tr>
<tr>
<td>C</td>
<td>320 m</td>
</tr>
<tr>
<td>D</td>
<td>405 m</td>
</tr>
<tr>
<td>E</td>
<td>640 m</td>
</tr>
</tbody>
</table>
### Alternatives

<table>
<thead>
<tr>
<th>OPTION</th>
<th>ALTERNATIVE</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>100 m</td>
</tr>
<tr>
<td>B</td>
<td>300 m</td>
</tr>
<tr>
<td>C</td>
<td>320 m</td>
</tr>
<tr>
<td>D</td>
<td>405 m</td>
</tr>
<tr>
<td>E</td>
<td>640 m</td>
</tr>
</tbody>
</table>

You selected C when answering this question
The contributor suggests C is the correct option

<table>
<thead>
<tr>
<th></th>
<th>RESPONSES</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1 (2.17%)</td>
</tr>
<tr>
<td>B</td>
<td>1 (2.17%)</td>
</tr>
<tr>
<td>C</td>
<td>26 (56.52%)</td>
</tr>
<tr>
<td>D</td>
<td>16 (34.78%)</td>
</tr>
<tr>
<td>E</td>
<td>2 (4.35%)</td>
</tr>
</tbody>
</table>

### Explanation

The following explanation has been provided relating to this question:

Calculate the average time taken

\[ \frac{8.8 + 9.5 + 8.7 + 9.2 + 8.8}{5} = 9.0 \text{ s} \]

The total time taken is 9.0 seconds which includes \( t_{bell} \) and \( t_{sound} \)
Consider the bell:

Height = \frac{1}{2} g t_{bell}^2 + v_0 t_{bell}, \text{ where } v_0 = 0 \text{ m/s}^{-1}

So, Height = \frac{1}{2} g t_{bell}^2 = \frac{1}{2} g (t_{total} - t_{sound})^2 = \frac{1}{2} \times 10 \times (9.0 - t_{sound})^2. \quad (1)

Now, think about the sound:

Height = V_{sound} t_{sound} = 320 t_{sound} \quad \cdots \quad (2)

Solve the equations (1) and (2)

(1) - (2) = 0

\frac{1}{2} \times 10 \times (9.0 - t_{sound})^2 - 320 t_{sound} = 0

t_{sound} = 1.0 \text{ s}

Height = V_{sound} t_{sound} = 1.0 \times 320 = 320 \text{ m}
Please rate this question:

Please rate this question as fairly and accurately as you can - your rating will help others to find questions of interest.

**Difficulty**

- Easy
- Medium
- Hard

**Quality**

- very poor
- poor
- fair
- good
- very good
- excellent

0 1 2 3 4 5

**Comment**

There are 25 comments written about this question.

<table>
<thead>
<tr>
<th>WHEN</th>
<th>COMMENT</th>
<th>AGREE WITH COMMENT</th>
<th>DISAGREE WITH COMMENT</th>
</tr>
</thead>
<tbody>
<tr>
<td>7:38pm, 20 Oct</td>
<td><strong>⭐⭐⭐⭐</strong> 1240 Nice problem, never done a question like this before. <strong>Author's reply:</strong> Thank you... £º</td>
<td>🌟</td>
<td>✗</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Bloom’s Taxonomy (Revised)


based on an APA adaptation of Anderson, L.W. & Krathwohl, D.R. (Eds.) (2001)
• 2007-Summer 2010
  – 45 institutions
  – 260 courses
  – 20661 students have contributed
  – 57324 questions have been written
  – 1527574 answers have been submitted
• Feb 2011
  – 77 institutions
  – 557 courses
  – 33757 students have contributed
  – 94207 questions have been written
  – 2308854 answers have been submitted
• Oct 2012
  – 308 institutions
  – 1905 courses
  – 94961 students have contributed
  – 379464 questions have been written
  – 8172405 answers have been submitted
II. Use at UoE
Use of the system at Edinburgh, 2010-present

- First-year, calculus-based introductory courses
  - Semester 1:
    - Newtonian mechanics
  - Semester 2:
    - Waves and modern physics
- Cohort:
  - Approx. 200-300 students
  - 75% male, 25% female
  - 50% majors, 50% non-majors
Pilot year (2010-11) – replace single handin

PeerWise was introduced in workshop sessions in Week 5

Students worked through structured example task and devised own Qs in groups.

All these resources are available online (see final slide)
An assessment was set for the end of Week 6:

Minimum requirements:

• Write one question
• Answer 5
• Comment on & rate 3

Contributed ~3% to course assessment
Rollout year 2011-12: 3 hand-ins replaced

Same requirements each time (w1, a5, r&c 3)

- Activity 1: intro as before
- Activity 2: focus on distracters
- Activity 3: integrating diff. sections of course

Contributed ~7% to course assessment
Screencasts

The following set of screencasts are provided courtesy of the Physics Education Research Group at the University of Edinburgh. Just select the screencast you would like to view from the list below:

Creating questions in PeerWise

This screencast illustrates the process of creating a new question - writing the question stem and alternatives, selecting the correct answer, providing an explanation, and tagging the question with relevant topics.

Editing questions in PeerWise

This screencast illustrates the process of making changes to an existing question, in response to feedback provided on the question.

Searching for questions on PeerWise

This screencast illustrates the process of searching for questions of interest on PeerWise, including filtering by topic, sorting by quality ratings and following question authors.

How to register and log into

This screencast, shown from the perspective of a student at the University of Edinburgh, illustrates the process for registering a new PeerWise account and logging in.
We were *deliberately* hands off.

- No moderation
- No corrections
- No interventions at all

But we did *observe*.....
Generally, students did

- Participate beyond minimum requirements
- Engage in community learning, correcting errors
- Create problems, not exercises
- Provide positive feedback
James Bond (mass 70kg) is trying to escape a building by abseiling out the window of an office. He is tethered via a light inextensible rope to a baddie he just knocked out inside the office. The rope passes around him over a rounded window ledge. The unconscious baddie is lying stationary on the office floor, with a static coefficient of friction of 0.7 between him and the floor.

**Alternatives**

<table>
<thead>
<tr>
<th>OPTION</th>
<th>ALTERNATIVE</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0.3</td>
</tr>
<tr>
<td>B</td>
<td>0.4</td>
</tr>
<tr>
<td>C</td>
<td>0.5</td>
</tr>
<tr>
<td>D</td>
<td>0.6</td>
</tr>
<tr>
<td>E</td>
<td>0.7</td>
</tr>
</tbody>
</table>

(The commissioner suggests C is the correct option.)

**Explanation**

Bond briefly tugs on the rope and begins accelerating vertically down the building at a constant $1 \text{m/s}^2$, simultaneously dragging the baddie horizontally across the office floor in the process. What is the coefficient of kinetic friction between the baddie and the floor?
Generally, students did not

• Contribute trivial or irrelevant questions
• Obviously plagiarise
• Participate much beyond assessment periods
• Didn’t all leave it to the last minute
Correlation with end of course outcomes

<table>
<thead>
<tr>
<th>Cohort</th>
<th>Number of students</th>
<th>Mean exam score*</th>
<th>Standard error</th>
<th>p value</th>
<th>Effect size</th>
</tr>
</thead>
<tbody>
<tr>
<td>1A (N=193)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HPA†</td>
<td>104</td>
<td>63.2</td>
<td>1.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LPA</td>
<td>89</td>
<td>53.6</td>
<td>1.6</td>
<td>&lt;0.001</td>
<td>0.29</td>
</tr>
<tr>
<td>1B (N=182)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HPA</td>
<td>94</td>
<td>61.9</td>
<td>1.8</td>
<td>&lt;0.001</td>
<td>0.36</td>
</tr>
<tr>
<td>LPA</td>
<td>88</td>
<td>46.8</td>
<td>2.4</td>
<td>&lt;0.001</td>
<td></td>
</tr>
</tbody>
</table>

* all scores expressed as percentages
† HPA / LPA denote higher / lower than median PeerWise activity
Quartiles
Q1 – top 25%
Q2 – upper middle
Q3 – lower middle
Q4 – bottom 25%

22 students did not take the FCI
Median Split for IgCA4

Error Bars: +/- 1 SE
Student-Generated Content for Learning (SGC4L)

Summary

One of the key attributes that undergraduate study of a subject seeks to develop is an advanced level of problem solving ability within the discipline. This is particularly true in, although not restricted to, science disciplines. Although deliberate practice can develop these skills, it has been argued that a deeper understanding can be achieved by having students pose, as well as answer, problems. In cognitive terms, it is far more demanding to generate both correct and incorrect reasoning and answers to a problem than merely attempting to find a solution.

PeerWise (http://peerwise.cs.auckland.ac.nz/) is a freely available web tool that provides an online framework to facilitate student creation of problems as well as including much of the social functionality that increasingly forms the cornerstone of online interactions. Using the tool, students can create assessment questions (in the form of multiple choice questions, with associated explanations), answer each other’s questions, rate and comment on questions, seek help from authors and follow their favourite question contributors. If embedded appropriately in course assessment design, use of the system offers tangible benefits to both students and staff, enabling valuable peer discussion, interaction and feedback outside timetabled class hours.
Results – Second Year Physics, University of Glasgow

<table>
<thead>
<tr>
<th>Cohort</th>
<th>Number of Students</th>
<th>Mean Exam Score *</th>
<th>Standard Error</th>
<th>p value</th>
<th>Effect Size **</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physics 2</td>
<td>(N=152)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HPA ***</td>
<td>78</td>
<td>58.1</td>
<td>1.9</td>
<td>&lt;0.001</td>
<td>0.45</td>
</tr>
<tr>
<td>LPA</td>
<td>74</td>
<td>38.0</td>
<td>2.7</td>
<td>&lt;0.001</td>
<td></td>
</tr>
</tbody>
</table>

* all scores expressed as percentages  
** Pearson’s r  
*** HPA/LPA denote higher/lower PeerWise activity
Results – Second Year Physics, University of Glasgow
III. Question quality
Comprehensive categorisation of >50% of repository for two successive academic years

Principal measures to define a ‘high quality question’

- cognitive level of question
- explanation quality
- other criteria
Cognitive level of question

<table>
<thead>
<tr>
<th>Category</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>Create (synthesize ideas)</td>
</tr>
<tr>
<td>5</td>
<td>Assess</td>
</tr>
<tr>
<td>4</td>
<td>Analyse (multi-step)</td>
</tr>
<tr>
<td>3</td>
<td>Apply (1-step calcs.)</td>
</tr>
<tr>
<td>2</td>
<td>Understand</td>
</tr>
<tr>
<td>1</td>
<td>Remember</td>
</tr>
</tbody>
</table>

Based on an APA adaptation of Anderson, L.W. & Krathwohl, D.R. (Eds.) (2001)
## Cognitive level of question

<table>
<thead>
<tr>
<th>Rating:</th>
<th>Description:</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Remember, Recognise or Recall OR just plugging in numbers</td>
</tr>
<tr>
<td>2</td>
<td>Understand, Interpret or predict (No calculation needed, such as understanding Newtons 3\textsuperscript{rd} law)</td>
</tr>
<tr>
<td>3</td>
<td>Apply, Implement or Calculate (1 step calculation)</td>
</tr>
<tr>
<td>4</td>
<td>Analyse, differentiate or organise (multi-step calculation, higher analysis)</td>
</tr>
<tr>
<td>5</td>
<td>Evaluate, Asses or Rank (Evaluating various options and assessing their validity)</td>
</tr>
<tr>
<td>6</td>
<td>Create, Combine or Produce (Asked to combine various areas of physics, need to get a structure right to solve whole problem)</td>
</tr>
</tbody>
</table>
Explanation

0 – Missing

1 – Inadequate
(e.g. wrong reasoning / answer, trivial, flippant, unhelpful)

2 – Minimal
(e.g. correct answer, but with insufficient explanation or justification, aspects may be unclear)

3 – Good/Detailed
(e.g. clear and sufficiently detailed exposition of correct method and answer)

4 – Excellent
(e.g. Describes physics thoroughly, remarks on plausibility of answer, use of appropriate diagrams, perhaps explains reasoning for distractors)
‘High quality’ question

1. At least 2/6 on cognitive level ("understand" and above)
2. At least 2/4 on explanation ("minimal" and above)
3. Clearly worded question (binary)
4. Feasible distractors
5. ‘Most likely’ correct (binary)
6. ‘Not obviously’ plagiarised (binary)
Categorisation process

- 2 raters: categorise ~35 questions
- Initial inter-rater reliability check: refine
- Categorise further 22 questions

- IRR determined using Cohen's Kappa.

- Agreement above 90% for taxonomic level and explanation.
Example category 1 question

A spectral line is emitted when an electron in a hydrogen atom undergoes a transition from n=5 to n=3 state. State which series this line belongs to.

**Alternatives**

<table>
<thead>
<tr>
<th>OPTION</th>
<th>ALTERNATIVE</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Lyman</td>
</tr>
<tr>
<td>B</td>
<td>Balmer</td>
</tr>
<tr>
<td>C</td>
<td>Paschen</td>
</tr>
<tr>
<td>D</td>
<td>None of the above</td>
</tr>
</tbody>
</table>
Example category 2 question

The diagram shows four electron energy levels in an atom. The transition of an electron from level 3 to level 1 as shown in the diagram produces a photon in the visible light range. Which transition is most likely to produce a photon in the ultraviolet range?

Alternatives

<table>
<thead>
<tr>
<th>OPTION</th>
<th>ALTERNATIVE</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Level 2 to level 1</td>
</tr>
<tr>
<td>B</td>
<td>Level 3 to level 2</td>
</tr>
<tr>
<td>C</td>
<td>Level 4 to level 1</td>
</tr>
<tr>
<td>D</td>
<td>Level 4 to level 3</td>
</tr>
</tbody>
</table>
Example category 3 question

The half-life of Po 210 is 138 days. How long does it take 18g of Po 210 to decay to only 2.25g?

**Alternatives**

<table>
<thead>
<tr>
<th>OPTION</th>
<th>ALTERNATIVE</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>552 days</td>
</tr>
<tr>
<td><strong>B</strong></td>
<td>414 days</td>
</tr>
<tr>
<td>C</td>
<td>1104 days</td>
</tr>
<tr>
<td>D</td>
<td>276 days</td>
</tr>
</tbody>
</table>
Example category 6 question

King Kong and Godzilla are slugging it out in downtown Tokyo, as they are prone to do on quiet Sunday evenings.

Kong quickly gains the upper hand and catches Godzilla by the tail, spinning him in a tilted circle of radius $r$ meters, in a clockwise direction, at an angle of $\alpha$ degrees to the horizontal. (fig. A)
A The block can't be at rest. In fact, it experiences a net external force $F_{\text{net}} = F + W + T$ where $F$ is the spring tension, $W$ is the weight, and $T$ is the friction force. Then, applying Newton's second law, we deduce the block is accelerated.

B You need to consider $F$ as a sum of two vectors $F_N$ and $F_F$. $F_N$ is the normal contact force and $F_F$ is the friction force.

C The horizontal component of $F$ (so $F_F$) is pointing to the left. But friction forces (in situations like this) are always opposing the direction of motion. Thus, the block is moving towards the right.

D The spring tension is in the opposite direction of motion (from the picture and C).

E Weight is constant. In addition, it is perpendicular to the direction of motion. It is not involved in the horizontal motion of the block.
I could not figure out how to do this but I figured it has to be either 600N or 700N as T has to be > 400N. Anyway after reading your explanation, I noticed something went wrong.

Quoted:

1) Mass

\[ m_1 \]
\[ G_{1x} - F_{f1} - T_1 = m_1 \cdot a \]
\[ R_1 = G_{1y} \]

\[ F_{f1} = \mu_k \cdot R_1 \]

If we work out on the equations we get:

\[ m_1 \cdot g \cdot \sin 60 - \mu_k \cdot m_1 \cdot g \cdot \cos 60 - T_1 = m_1 \cdot a \]

As you can see you equated \( G_{1x} = m_1 \cdot g \cdot \sin 60 \) when it meant to be \( m_1 \cdot g \cdot \cos 60 \). I think you got that one switched around with the \( R_1 = G_{1y} \). Same thing happened with equation 2.

I might also be wrong so do have a look at it.

(by: duckula)

**Author's reply**

No, it is correct. \( G_{1x} = m_1 \cdot g \cdot \sin 60 \). Think what happens when the angle is 0 (the mass is on the horizontal). \( G_{1x} = m_1 \cdot g \cdot \sin 0 = 0 \).

Here is a sketch, to make things clear.
Results: Physics 1A 2010 and 2011

2 successive years of the same course (N=150, 350)

- ‘High quality’ questions: 78%, 79%
- Over 90% (most likely) correct, and 3/5 of those wrong were identified by students.
- 69% (2010) and 55% (2011) rated 3 or 4 for explanations
- Only 2% (2010) and 4% (2011) rated 1/6 for taxonomic level.
Results: Question level Physics 1A 2010 and 2011

Taxonomy Ratings

<table>
<thead>
<tr>
<th>Rating</th>
<th>Number of Answers</th>
<th>2010-11</th>
<th>2011-12</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Literature


- 107 Year 2 biochem students
- 56 / 35 / 9 % of questions in lowest 3 levels.


“9,713 assessment items submitted by 50 instructors in the United States reported that 93% of the questions asked on examinations in introductory biology courses were at the lowest two levels of the revised Bloom’s taxonomy”
Zheng et al. (2008), *Science* 319, pp414-415
Results: Explanation Physics 1A 2010 and 2011
Results: Question level Physics 1A / 1B 2011

First semester  N = 350
Second semester N = 252
Summary

- High general standard of student-generated questions
- Relatively few basic knowledge questions
- Vast majority of questions require at least application
- Some questions at highest cognitive levels
- Appears not to be course or subject specific
- We hypothesise scaffolding activities may promote high level cognitive engagement
IV. Community, further research
Further work

- Controlled experiment for influence of scaffolding?

3 groups:
- control (no intervention)
- partial (tool, no scaffolding)
- full (tool, scaffolding)
Further work

- Other correlations:
  - Who answers what? (social network analysis)
  - What’s the role / impact (if any) of comments?
  - Question quality $\leftrightarrow$ academic ability?

- Crowd-sourced assessments? (appropriately validated)

- Multi institution course space?
What does a typical PeerWise course look like?

by Paul Denny

October 12, 2012 in Uncategorized

If you have ever wondered whether your class is too small (or too big) to use a tool like PeerWise, you may be interested in the following data. To get a sense for both the typical size of a class on PeerWise, and the typical number of contributions made by students in each class, data from the last 1000 courses was examined.

While there are many examples of very large classes (>300 students), and even a few extremely large ones (>800 students), the majority of classes have fewer than 50 students. The breakdown is given in the chart below.

**Typical class sizes - a summary of 1000 PeerWise courses**
Acknowledgements

Ross Galloway
Judy Hardy
Karon McBride
Alison Kay
Keith Brunton
Jonathan Riise
Danny Homer

Chemistry – Peter Kirsop
Biology – Heather McQueen

Physics – Morag Casey
Comp Sci – Paul Denny
Resources

Community:
http://www.PeerWise-Community.org

JISC-funded multi institution study:
https://www.wiki.ed.ac.uk/display/SGC4L/Home

UoE Physics Pilot Study:
AIP Conf. Proc. 1413, 359 http://dx.doi.org/10.1063/1.3680069

UoE Physics scaffolding resources
http://www2.ph.ed.ac.uk/elearning/projects/peerwise/
a place of mind
THE UNIVERSITY OF BRITISH COLUMBIA