Use of Flawed MCQ's in Final Exams

Which of the following is the largest?
A: peanut
B: An Elephant
C: The Moon
D: A Kettle

Yan Jin
WS Chan
Simon Yip
Kumta Shekhar
• Tests are a key component of assessing students’ knowledge.

• Since test grades affect students’ educational outcomes and subsequent career paths, test items should be well constructed.

• If tests are not well constructed, assessments of student performance may be invalid.

• Good item construction is critical to accurate assessment
What is a Flawed MCQ

• Guidelines for MCQ construction are clearly documented in the many books and articles that have been written on this topic from a variety of disciplines.

• Item-writing flaws (IWFs) are violations of these accepted item-writing guidelines which can affect student performance on MCQs, making the question either easier or more difficult to answer.
Item Writing Flaws

(there are 19 categories of IWF's - here are some common ones)

- ambiguous or unclear information in the stem
- negatively worded stems
- implausible distracters
- unnecessary or gratuitous information in the stem
- more than one or no correct answer
- the longest option is correct
- logical clues in the stem
- word repeats in the stem and correct answer
- grammatical clues associated with sentence completions
- convergence clues
Example

• About Diabetes:

• Which of the following is NOT true.

• A: Hb A1C is best for monitoring glycaemic control.

• B: Limb Salvage is possible in most cases of lower-limb ulceration if glycaemic control is good.

• C: Risk of Cardiovascular mortality is related to obesity not glycaemic control.
Another one....

- The actions of Thyroxine include all except........
- Which of the following is NOT an effect of the drug xxxxxxxxx

What is NOT likely to pop-out from this magician's hat

Difficulty related to construct rather than content
Impact of Flawed Items

- Students confused
- Affects pass-fail standards
- May affect test validity
- Cannot be ethically & legally defended

CUHK Exams

<table>
<thead>
<tr>
<th>Year</th>
<th>No of MCQ's</th>
<th>Flawed</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>2006-07</td>
<td>150</td>
<td>24</td>
<td>16%</td>
</tr>
<tr>
<td>2007-08</td>
<td>150</td>
<td>41</td>
<td>27%</td>
</tr>
<tr>
<td>2007-09</td>
<td>150</td>
<td>22</td>
<td>15%</td>
</tr>
<tr>
<td>2007-08</td>
<td>150</td>
<td>14</td>
<td>9%</td>
</tr>
<tr>
<td>2008-09</td>
<td>150</td>
<td>19</td>
<td>13%</td>
</tr>
</tbody>
</table>

Mean 16%
# Impact on Pass-Fail Stats

<table>
<thead>
<tr>
<th>Passing Standard (%)</th>
<th>Passing Rates as Percentage</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>With flawed items</td>
<td>Without flawed Items</td>
</tr>
<tr>
<td>78</td>
<td>4.6</td>
<td>5.8</td>
</tr>
<tr>
<td>70</td>
<td>30.4</td>
<td>39.8</td>
</tr>
<tr>
<td>64</td>
<td>61.6</td>
<td>68.1</td>
</tr>
<tr>
<td>58</td>
<td>84.1</td>
<td>90.6</td>
</tr>
<tr>
<td>54</td>
<td>92.8</td>
<td>94.2</td>
</tr>
</tbody>
</table>

Affects 9 students
## Impact on Pass-Fail Stats

<table>
<thead>
<tr>
<th>Passing Standard (%)</th>
<th>Passing Rates as Percentage</th>
<th>Without flawed Items</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>With flawed items</td>
<td></td>
<td></td>
</tr>
<tr>
<td>76</td>
<td>2.2</td>
<td>5.2</td>
<td>3.0</td>
</tr>
<tr>
<td>70</td>
<td>17.9</td>
<td>26.8</td>
<td>8.9</td>
</tr>
<tr>
<td>64</td>
<td>48.5</td>
<td>59.0</td>
<td>10.5</td>
</tr>
<tr>
<td>58</td>
<td>87.3</td>
<td>91.8</td>
<td>4.5</td>
</tr>
<tr>
<td>52</td>
<td>97.8</td>
<td>99.2</td>
<td>1.4</td>
</tr>
</tbody>
</table>

Affects 15 students
How may we avoid IWF's:

- Use a Test Committee
  - single writer is more error prone
- use a blue-print
  - sample broad, set appropriate cognitive level
- set standards
  - use items of appropriate difficulty
- Review the Exam paper before and after the exam
  - identify IWF's and Content Validity
- Use MCQ' banks
  - IDEAL vets flawed items
- Attend an ITEM-writing workshop
- Use OES-TLRC staff with expertise and experience