Scratch and Win!
Immediate feedback in physics multiple-choice quizzes and exams

Jeffrey L. Hutter
Carol E. Jones
Margaret Campbell-Brown
Dept. of Physics & Astronomy
The University of Western Ontario

Western

Summary

- Multiple-choice assessments
  - pros & cons
- The Immediate Feedback Assessment Technique (IF-AT)
  - the technique, pros & cons, previous studies
- Our study
  - assessing the IF-AT in a large first-year physics class
- Conclusions
  - how well does it work?
  - recommendations for prospective IF-AT users
  - what comes next?

Multiple Choice Assessments

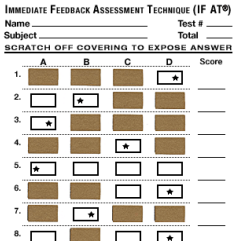
- Advantages of multiple-choice (MC) exams:
  - ability to quickly test a wide range of concepts
  - availability of test banks
  - easily graded
  - multiple versions can make cheating more difficult, and easy to detect
- Disadvantages:
  - difficult to construct an exam that tests concepts rather than memorization
    - makes it undesirable to provide solutions or to post previous exams
  - most students never review their wrong answers

Effectiveness of MC Assessments

- A "good" multiple-choice exam *can* be constructed ...
  - well-written MC exams can effectively test student understanding
    - M.G. Simkin and W.L. Kuechler, *Decision Sciences Journal of Innovative Education* **3**, 73 (2005).
- ... but feedback is crucial to student learning.
  - and early feedback is more effective than delayed feedback
    - R.L. Bangert-Drowns, et al., *Review of Educational Research* **61**(2), 213 (1991).
    - R.E. Dihoff, et al., *The Psychological Record* **54**, 207 (2004).
- The problem:
  - how do you provide feedback for MC exams while preserving exam security?

The Immediate Feedback Assessment Technique

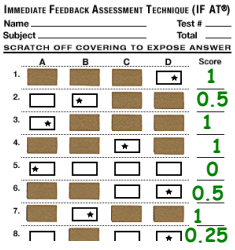
- The IF-AT technique:
  - a way of implementing feedback in MC exams
    - M.L. Epstein, et al. *Psychological Reports* **88**, 889 (2001).
  - students answer MC questions by uncovering an opaque waxy coating on a special answer card
  - if a star is uncovered, the answer is correct
  - if the answer is wrong, students can review their reasoning and try again



Epstein Educational Resources
www.epsteineducation.com

IF-AT Advantages

- Advantages of the IF-AT:
  - immediate feedback
    - the exam itself becomes a learning experience
    - no need to post answer keys
  - partial credit in a multiple-choice exam
    - popular with students!
  - no need to double-check answers
    - students know their scores before leaving the exam



IF-AT Drawbacks

- Disadvantages of the IF-AT:
  - more work to set up the exam
    - instructors must arrange answers to match a limited variety of IF-AT cards
  - answer keys cannot be changed
    - instructors must get it right the first time
  - not all answer types and orders are appropriate
    - e.g., "none of the above" and "all of the above" are awkward
  - not currently machine-readable
    - must be hand graded
    - harder to detect cheating
  - expense
    - ~$500 / 2000 cards
  - immediate feedback
    - can be discouraging for some students

Effectiveness of the IF-AT

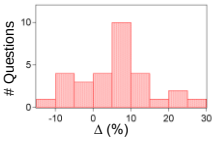
- Literature findings:
  - the IF-AT improves learning/retention of knowledge
    - M.L. Epstein, et al., *The Psychological Record* **52**, 187 (2002).
  - the IF-AT is popular with students
    - D. DiBattista et al., *Teaching in Higher Education* **9**(1), 17 (2004).
    - D. DiBattista and L. Gosse, *The Journal of Experimental Education* **74**(4), 311 (2006).
- But almost **no** data for science courses, none for physics.
- The question to be addressed:
  - Is the IF-AT worth the trouble in a quantitative Science course?

The Study

- Western Teaching Support Centre Grant of ~$2000 to:
  - purchase IF-AT cards
  - use the IF-AT in tutorials and exams for Physics 1024
    - standard calculus-based course in first-year physics
  - assess the IF-AT & compare with a previous (non-IF-AT) class
- Methods:
  - ten biweekly tutorial quizzes (8 MC questions + 1 problem)
    - using the IF-AT
    - 3 different tutorial sections with different problems
  - two midterm exams (24 MC + 4 problems)
    - using the IF-AT
  - one final exam (28 MC + 5 problems)
    - using Scantrons
  - collect data from consenting students, anonymize
  - conduct exit survey

Overall Results

- Question 1: Does exposure to a question in a quiz improve performance on the same question in an exam?
  - we chose 30 questions to appear in both IF-AT quizzes and exams
    - in each case, only some tutorial sections saw the question
  - compared the exam scores between groups who had been exposed to a question, and those who had not
    - Δ = % difference in average performance
  - on average, groups who had previously seen a question did better than groups who had not
    - $\langle\Delta\rangle = (6.2 \pm 1.7)$ %



MC Questions with a large Δ

14. Consider a sky surfer falling through air, before reaching her terminal speed because of air resistance. As the speed of the sky surfer increases, the magnitude of her acceleration:

(a) decreases until it reaches zero
(b) decreases until it reaches a constant nonzero value
(c) increases
(d) remains constant
(e) not enough information is given to answer this question

Midterm 1
Δ = 75.0 – 48.8
= 26.2

11. Neglecting air resistance, a 1.0 kg projectile has an escape velocity of about 11 km/s at the surface of the Earth. The corresponding escape velocity for a 2.0 kg projectile is:

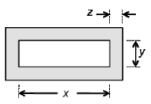
A. 22 km/s
B. 10 km/s
C. 11 km/s
D. 7.1 km/s
E. 5.5 km/s

Midterm 2
Δ = 71.4 – 47.6
= 26.2

MC Questions with a large Δ

1. The figure shows a rectangular brass plate at 0 °C in which there is cut a rectangular hole of dimensions x x y, as indicated. If the temperature of the plate is raised to 150 °C:

A. both x and y will decrease
B. both x and y will increase
C. the changes in x and y depend on the dimension z
D. x will decrease and y will increase
E. x will increase and y will decrease



Final
Δ = 88.4 – 71.3
= 17.2

- Common themes for large Δ:
  - previous exposure has the most benefit for simple questions that students are likely to get wrong the first time
    - helps to clear up misconceptions

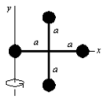
MC Questions with a modest Δ

6. A bullet shot horizontally from a gun:

(a) strikes the ground much later than one dropped vertically from the same point at the same distance
(b) never strikes the ground
(c) strikes the ground at approximately the same time as one dropped vertically from the same point at the same instant
(d) travels in a straight line
(e) strikes the ground much sooner than one dropped from the same point at the same instant

Midterm 1
Δ = 95.4 – 92.5
= 2.9

8. Four identical particles, each with mass m, are arranged in the x,y plane as shown. They are connected by light sticks to form a rigid body. If $m = 2.0$ kg and $a = 1.0$ m, the rotational inertia of this array about the y axis is:



A. 12 kg•m²
B. 4.8 kg•m²
C. 9.6 kg•m²
D. 6.0 kg•m²
E. 16 kg•m²

Midterm 2
Δ = 50.9 – 43.2
= 7.7

MC Questions with a modest Δ

18. A non-viscous incompressible fluid is pumped steadily into the narrow end of a long tapered pipe and emerges from the wide end. The pressure at the input is greater than at the output. A possible explanation is:

A. the fluid is flowing uphill
B. the fluid speed is the same at the two ends
C. the fluid is flowing horizontally
D. the fluid is flowing downhill
E. the fluid speed increases from input to output

Midterm 2
Δ = 60.4 – 52.9
= 7.4

- Common themes for modest Δ:
  - easy questions
  - questions that require application of concepts
    - both conceptual and computational questions

MC Questions with a *negative* Δ

5. The rate of heat flow by conduction through a slab does NOT depend upon the:

A. temperature difference between opposite faces of the slab
B. specific heat of the slab
C. thermal conductivity of the slab
D. cross-sectional area of the slab
E. slab thickness

Final
Δ = 61.7 – 71.2
= –9.5

21. The current is from left to right in the conductor shown. The magnetic field is into the page, and point S is at a higher electrical potential than point T. The charge carriers are:

A. negative
B. neutral
C. positive
D. moving near the speed of light
E. absent

Final
Δ = 46.8 – 53.0
= –6.2

Observations on Question Types

- The gain resulting from previous exposure to questions depends on the type of question:
  - easy questions ⇒ small Δ
    - both groups do well
  - difficult conceptual questions ⇒ small Δ
    - prior exposure helps, but Δ typically < 10%
  - difficult calculations ⇒ small Δ
    - prior exposure helps, but Δ typically < 10%
  - simple, but "non-intuitive" conceptual questions ⇒ large Δ
    - students retain the answer

- What type of questions result in a *negative* Δ?
  - mostly ones where later concepts might confuse students
  - probably the result of sample bias, though
    - the most negative values of Δ where achieved by a single tutorial section

A closer look

18. A non-viscous incompressible fluid is pumped steadily into the narrow end of a long tapered pipe and emerges from the wide end. The pressure at the input is greater than at the output. A possible explanation is:

A. the fluid is flowing uphill
B. the fluid speed is the same at the two ends
C. the fluid is flowing horizontally
D. the fluid is flowing downhill
E. the fluid speed increases from input to output

	Quiz Result	Exam Result (w/ previous exposure)	Exam Result (no exposure)
% 1s	33		
% 2s	22		
% 3s	16		
% 4s	9		
% 5s	21		

A closer look

11. Neglecting air resistance, a 1.0 kg projectile has an escape velocity of about 11 km/s at the surface of the Earth. The corresponding escape velocity for a 2.0 kg projectile is:

A. 22 km/s
B. 10 km/s
C. 11 km/s
D. 7.1 km/s
E. 5.5 km/s

	Quiz Result	Exam Result (w/ previous exposure)	Exam Result (no exposure)
% 1s			48
% 2s			21
% 3s			11
% 4s			12
% 5s			9

Observations on Question Types

- Question 2: Does the IF-AT work better than other feedback?
  - three of the final exam questions considered were also posed in both a non-IF-IT quiz and exam in a previous year

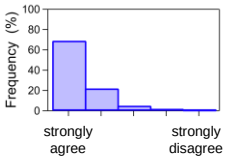
Question	IF-AT Δ (%)	non-IF-AT Δ (%)
4	5.4	
20	10.2	
21	–6.2	

  - apparently *conventional* multiple-choice (without feedback) works as well as the IF-AT!?
  - but 2 of the 3 questions in the IF-AT group were done by tutorial section 005...

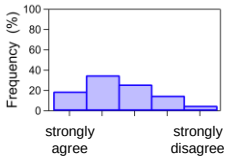
Student Acceptance

- Question 3: What do the students think?
  - Conducted a 15-question year-end survey to query student opinion.
    - 5-point responses ranging from "strongly agree" to "strongly disagree"

I preferred the IF-AT to Scantrons.

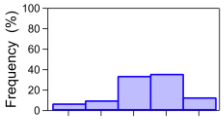


The I-FAT helped me retain knowledge.

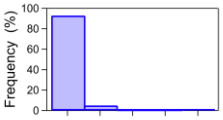


Student Acceptance

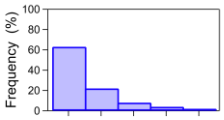
The IF-AT allowed me to complete the exam in less time.



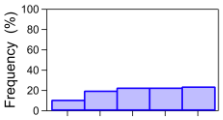
I liked being able to get partial credit.



I liked knowing my score when I completed the exam.



I found the IF-AT to be stressful.



Measuring the Impact

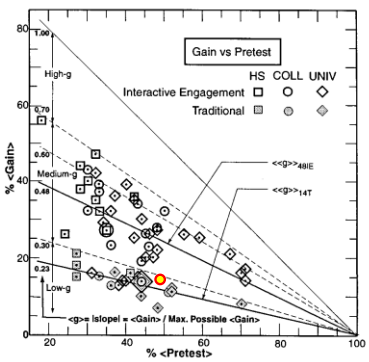
- One possible advantage of the IF-AT is that it makes exams interactive, teaching students to reason through problems
- Question 4: Do students trained using the IF-AT end up with a better knowledge of physics?
- We test general ability using the Force Concept Inventory (FCI) as a pre- and post-test.
  - the FCI is a well-known quiz testing basic force concepts
- FCI results have been studied for a variety of teaching method.
  - one famous result is that "interactive engagement" methods (e.g., peer instruction) result in much higher performance gains than traditional lecturing
    - R.R. Hake, *American Journal of Physics* **66**, 64 (1998).

Example FCI Question

- A large truck collides head-on with a small compact car. During the collision:

A) the truck exerts a greater amount of force on the car than the car exerts on the truck.
B) the car exerts a greater amount of force on the truck than the truck exerts on the car.
C) neither exerts a force on the other, the car gets smashed simply because it gets in the way of the truck.
D) the truck exerts a force on the car but the car does not exert a force on the truck.
E) the truck exerts the same amount of force on the car as the car exerts on the truck

Measuring the Impact



- Percentage gain in FCI scores plotted vs. initial score.
- Observations:
  - actual/potential gain seems to be a useful metric
  - interactive teaching methods offer a clear benefit
- No obvious benefit for IF-AT with traditional lectures.

Conclusions & Next Steps

- Students seem to be uniformly in favour of the IF-AT
- There seems to be little obvious advantage
  - some indication that traditional Scantrons are as effective
  - no obvious benefit to conceptual understanding of the material
  - ...*but* this is based on few comparisons
- Future analysis
  - more, and harder, questions
  - correlation with the problem-solving portion of the exams
  - retention vs. time
    - results from a six-week summer version of Physics 1024 showed similar trends
- *Why* is there no obvious advantage to the IF-AT in physics education when it has been proven to work in other fields?