Engaging reflective thinking during exam-like situations: Slowing students down on short-answer questions increases performance.

Jared Stang, Science Teaching and Learning Fellow

Joss Ives, Instructor I

Department of Physics and Astronomy, UBC

Carl Wieman Science Education Initiative

jared@phas.ubc.ca
@StangJared

joss@phas.ubc.ca
@jossives
Motivation:
Which (possibly extraneous) factors impact student performance on exams or in exam-like situations?
Background: Automatic cognitive processes play a role in student answers

- Dual processing theory describes two modes of thinking:

<table>
<thead>
<tr>
<th>System 1</th>
<th>System 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unconscious, intuitive judgements</td>
<td>Deliberate, conscious mental effort</td>
</tr>
</tbody>
</table>

Background: Automatic cognitive processes play a role in student answers

- Students process height (of A vs B) faster than slope
Background: Automatic cognitive processes play a role in student answers

- Students process height (of A vs B) faster than slope
- Imposing a delay increased performance (47% to 70%)

The Cognitive Reflection Test (CRT) measures tendency to engage system 2

- 3-item test
- Each item has an incorrect intuitive response

The Cognitive Reflection Test (CRT) measures tendency to engage system 2

- 3-item test
- Each item has an incorrect intuitive response

E.g:

A bat and a ball cost $1.10 in total. The bat costs $1.00 more than the ball. How much does the ball cost?
Motivation:
Which (possibly extraneous) factors impact student performance on exams or in exam-like situations?
Motivation:
Which (possibly extraneous) factors impact student performance on exams or in exam-like situations?
Does this correlate with the CRT?
Extraneous example: Question pairs

**Question Pair 1:** Two friends, Abbie and Bob, are using a setup consisting of two identical springs and two identical masses to investigate the behaviour of masses oscillating up and down on springs. They both pull their masses down, but Bob pulls his down further than Abbie.

**Q1a:** If they both release their masses at the exact same time, which mass will oscillate up and then back down and return to that person’s hand in the shortest time?

A) Abbie’s mass will return to her hand in the shortest time
B) Bob’s mass will return to his hand in the shortest time
C) Each mass will take the same amount of time to return to the hand of the person that released it

**Q1b:** If they both release their masses at the exact same time, which mass will have the highest maximum speed while it travels up and then back down again?

A) Abbie’s mass will have the highest maximum speed
B) Bob’s mass will have the highest maximum speed
C) Both masses will have the same maximum speed
Hypothesis: Student response can be influenced by asking about related concepts together ("Treatment") versus in isolation ("Control").

Used on 9 question pairs on an end-of-term diagnostic

First-year physics, N = 650 students
Question Pair 1: Two friends, Abbie and Bob, are using a setup consisting of two identical springs and two identical masses to investigate the behaviour of masses oscillating up and down on springs. They both pull their masses down, but Bob pulls his down further than Abbie.

Q1a: If they both release their masses at the exact same time, which mass will oscillate up and then back down and return to that person’s hand in the shortest time?

A) Abbie’s mass will return to her hand in the shortest time  
B) Bob’s mass will return to his hand in the shortest time  
C) Each mass will take the same amount of time to return to the hand of the person that released it

Q1b: If they both release their masses at the exact same time, which mass will have the highest maximum speed while it travels up and then back down again?

A) Abbie’s mass will have the highest maximum speed  
B) Bob’s mass will have the highest maximum speed  
C) Both masses will have the same maximum speed
Extraneous example: Question pairs

**Question Pair 1**: Two friends, Abbie and Bob, are using a setup consisting of two identical springs and two identical masses the behaviour of masses oscillating up and down on springs. They both pull their masses down, but Bob pulls his down further than Abbie.

**Q1a**: If they both release their masses at the exact same time, which mass will oscillate up and then back down and return to that person’s hand in the shortest time?

A) Abbie’s mass will return to her hand in the shortest time
B) Bob’s mass will return to his hand in the shortest time
C) Each mass will take the same amount of time to return to the hand of the person that released it

**Q1b**: If they both release their masses at the exact same time, which mass will have the highest maximum speed while it travels up and then back down again?

A) Abbie’s mass will have the highest maximum speed
B) Bob’s mass will have the highest maximum speed
C) Both masses will have the same maximum speed
**Question Pair 1:** Two friends, Abbie and Bob, are using a setup consisting of two identical springs and two identical masses to investigate the behaviour of masses oscillating up and down on springs. They both pull their masses down, but Bob pulls his down further than Abbie.

**Q1a:** If they both release their masses at the exact same time, which mass will oscillate up and then back down and return to that person’s hand in the shortest time?

A) Abbie’s mass will return to her hand in the shortest time  
B) Bob’s mass will return to his hand in the shortest time  
C) Each mass will take the same amount of time to return to the hand of the person that released it

**Q1b:** If they both release their masses at the exact same time, which mass will have the highest maximum speed while it travels up and then back down again?

A) Abbie’s mass will have the highest maximum speed  
B) Bob’s mass will have the highest maximum speed  
C) Both masses will have the same maximum speed
Results: A statistically significant improvement in question performance when pairs asked together (“Treatment”)

<table>
<thead>
<tr>
<th></th>
<th>p value</th>
<th>Odds ratio (95% CI)</th>
<th>Effect size (Cohen’s d)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fisher’s Exact Test</td>
<td>&lt;.001</td>
<td>1.43 (1.28, 1.60)</td>
<td>0.19 (0.14, 0.26)</td>
</tr>
<tr>
<td>Logistic Regression</td>
<td>&lt;.001</td>
<td>1.53 (1.36, 1.72)</td>
<td>0.23 (0.17, 0.30)</td>
</tr>
</tbody>
</table>
Results: Performance correlates with CRT score

Results: The “Treatment” does not appear to impact students with different CRT scores differently.

Comparisons made using Fisher’s Exact Test; Cohen’s $d$ reported for effect size with 95% Confidence intervals.
Experiment 2: “Explain your answer”

• Idea: Slow students down by including an “Explain your answer” box with short-answer questions (“Treatment”):

[0.5 points] Explain your reasoning in 1-2 sentences:

• Used on two questions on each of three exams
• First-year physics, N = 650 students
Results: A statistically significant improvement in question performance when asked to explain (“Treatment”)
Summary: Extraneous factors impact student performance

• Both **Question pairs** and **Explain your answer** impacted student performance, *without changing the question*

• Competing relevant and irrelevant dimensions may be mediated by automatic processes (but CRT did not interact with Treatment)

**Broader questions:** What does developing expertise look like, and how do we assess it?