4-2016

Probing Question Order Effect in Chemistry Concept Inventories

Molly Undersander  
*University of Nebraska - Lincoln*, molly.undersander@gmail.com

Travis J. Lund  
*Oregon Institute of Technology*, travis.lund@oit.edu

Laurie S. Langdon  
*University of Colorado Boulder*, laurie.langdon@colorado.edu

Marilyne Stains  
*University of Nebraska - Lincoln*, mstains2@unl.edu

Follow this and additional works at: [http://digitalcommons.unl.edu/ucareresearch](http://digitalcommons.unl.edu/ucareresearch)

Part of the Science and Mathematics Education Commons

[http://digitalcommons.unl.edu/ucareresearch/1](http://digitalcommons.unl.edu/ucareresearch/1)
Probing Question Order Effect in Chemistry Concept Inventories

Molly Undersander¹ (molly.undersander@gmail.com), Travis J. Lund², Laurie Langdon³, Marilyne Stains¹

¹1. Department of Chemistry, University of Nebraska – Lincoln ²2. Department of Natural Sciences, Oregon Institute of Technology ³3. Department of Chemistry and Biochemistry, University of Colorado – Boulder

What is Question Order Effect?

Test PV
1. Question 1
2. Pictorial
3. Question 3
4. Verbal
5. Question 5

Test VP
1. Question 1
2. Verbal
3. Question 3
4. Pictorial
5. Question 5

Teachers often randomize test questions and create multiple versions of tests to prevent cheating.

Current literature across various subjects is split on whether or not this gives students taking a certain test version an unfair advantage.¹²

The goal of this project is to test whether the question order effect is present in a chemistry concept inventory. Many studies have been done regarding content order and difficulty order, but we want to test the effect of pictorial versus verbal question order.

A similar study was performed at a western institution to compare results between institutions.

Research Question

• How does question order affect student performance on conceptually isomorphic questions when presented with pictorial and verbal versions of the questions?

Methods and Participants

• A 20 question concept inventory about acids and bases was given to all sections of General Chemistry II (GCII) and Organic Chemistry I (OCI) at the beginning and end of the semester for two semesters.

• Data was only kept if students answered with a proper level of effort. We only kept students who self-reported a moderate effort (1, 2, or 3 out of a 4 pt scale) and high effort (1 or 2 out of a 4 pt scale).

• 768 pre and post survey responses were collected from GCII and OCI. After cleaning the data for effort levels and whether or not the students used resources, we were left with the following sample size:

<table>
<thead>
<tr>
<th></th>
<th>GCII Pre</th>
<th>OCI Pre</th>
<th>GCII Post</th>
<th>OCI Post</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test Version</td>
<td>PV</td>
<td>VP</td>
<td>PV</td>
<td>VP</td>
</tr>
<tr>
<td>Question</td>
<td>Significance</td>
<td>effect size</td>
<td>Significance</td>
<td>effect size</td>
</tr>
<tr>
<td>P</td>
<td>15.7%</td>
<td>20.8%</td>
<td>0.267</td>
<td>0.67</td>
</tr>
<tr>
<td>V</td>
<td>30.6%</td>
<td>32.6%</td>
<td>0.715</td>
<td>0.22</td>
</tr>
</tbody>
</table>

• We observed a question order effect only on question V for students with high effort. The western institution found similarly insignificant results. The only question that showed significance was question Post P in the post populations.

• Due to this both institutions treated the data by separating each of the populations by gender as well as by students’ scores on the first bipartite homogenous questions to see if any of these factors played a role. No question order effect was observed.

Results – Concept Inventory

Moderate Effort % Correct

<table>
<thead>
<tr>
<th>Question</th>
<th>Test Version</th>
<th>Significance</th>
<th>effect size</th>
</tr>
</thead>
<tbody>
<tr>
<td>P</td>
<td>16.9%</td>
<td>19.7%</td>
<td>0.692</td>
</tr>
<tr>
<td>V</td>
<td>30.5%</td>
<td>36.4%</td>
<td>0.489</td>
</tr>
</tbody>
</table>

High Effort % Correct

<table>
<thead>
<tr>
<th>Question</th>
<th>Test Version</th>
<th>Significance</th>
<th>effect size</th>
</tr>
</thead>
<tbody>
<tr>
<td>P</td>
<td>16.9%</td>
<td>19.7%</td>
<td>0.692</td>
</tr>
<tr>
<td>V</td>
<td>30.5%</td>
<td>36.4%</td>
<td>0.489</td>
</tr>
</tbody>
</table>

Discussion/Conclusion

• The Concept Inventory data demonstrates that a question order effect does not exist among any of the populations.

Next Step

• We are now probing into question order effect in geoscience concept inventories using the same methods.

• So far, preliminary analysis shows similar results with significance in only one question (V) from the Moderate Effort Pre population.

• Example of pictorial question:

The maps below show the surface of the Earth as viewed from the sky.

Which map best illustrates where earthquake epicenters, marked with an X, would be located?

Future Work

• We will analyze students’ misconceptions in this inventory through item analysis.

• We also plan to look at how answer choices evolve from pre to post and across expertise level.

References


Acknowledgements

• UNL UCARE – Undergraduate Research Grant
• Stains Research Group

The University of Nebraska-Lincoln is an equal opportunity educator and employer.