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
1. Department of Chemistry, University of Nebraska – Lincoln 2. Department of Natural Sciences, Oregon Institute of Technology 3. 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. Question 3
3. Pictorial
4. Verbal
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.¹,²

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.

Method 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>PV</td>
<td>134</td>
<td>82</td>
<td>101</td>
<td>69</td>
</tr>
<tr>
<td>VP</td>
<td>144</td>
<td>81</td>
<td>101</td>
<td>66</td>
</tr>
</tbody>
</table>

• Selective, semi-structured interviews were conducted at the end of each semester. A total of 19 students were interviewed. (7 from GCII, 12 from OCI)

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.

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.

Results – Concept Inventory

<table>
<thead>
<tr>
<th>Instruments</th>
<th>GCII Pre</th>
<th>OCI Pre</th>
<th>GCII Post</th>
<th>OCI Post</th>
</tr>
</thead>
<tbody>
<tr>
<td>PV</td>
<td>15.7%</td>
<td>20.8%</td>
<td>11.0%</td>
<td>18.2%</td>
</tr>
<tr>
<td>VP</td>
<td>30.6%</td>
<td>32.6%</td>
<td>37.0%</td>
<td>38.2%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Question</th>
<th>Test Version</th>
<th>Significance</th>
<th>effect size</th>
<th>Question</th>
<th>Test Version</th>
<th>Significance</th>
<th>effect size</th>
</tr>
</thead>
<tbody>
<tr>
<td>P</td>
<td>15.7%</td>
<td>20.8%</td>
<td>0.267</td>
<td>P</td>
<td>15.7%</td>
<td>20.8%</td>
<td>0.267</td>
</tr>
<tr>
<td>V</td>
<td>30.6%</td>
<td>32.6%</td>
<td>0.715</td>
<td>V</td>
<td>30.6%</td>
<td>32.6%</td>
<td>0.715</td>
</tr>
</tbody>
</table>

Results – Interviews

<table>
<thead>
<tr>
<th>Question</th>
<th>Answer</th>
<th>GCII PV</th>
<th>OCI PV</th>
<th>GCII VP</th>
<th>OCI VP</th>
<th>Example of quote providing justification for choice</th>
</tr>
</thead>
<tbody>
<tr>
<td>Did you prefer seeing the verbal or pictorial question first?</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>&quot;I thought it [P] kind of helped to visualize the dissociated mess because you can tell the stronger ones and they tell you that’s dissociated and you know it’s a weak acid. So yeah that kind of helped me to see [P] first.&quot;</td>
</tr>
</tbody>
</table>
| Did the first question help with following questions? | 2 | 1 | 1 | 3 | 0 | "I did think that [V] influenced my answer because if I wasn’t 100% certain on the behavior of the strong and weak acid, I leaned back on my answer for [V] to answer [P], so by choosing an answer here in [V], I carried that information forward to [P]."

References


Acknowledgements

• UNL UCARE – Undergraduate Research Grant
• M. Stains Research Group