Using Computerized Lexical Analysis of Student Writing to Facilitate Just-in-Time Teaching in Large-Enrollment Biology Courses

Luanna Prevost, Kevin Haudek, Emily Norton Henry, Matthew Berry, Mark Urban-Lurain

Automated Analysis of Constructed Response (AACR) Research Group
www.msu.edu/~aacr
Assessment to Reveal Student Thinking

- Large enrollment courses prohibit the use of constructed response assessments
**Constructed Response (CR) Assessments**

- Allow students to represent their understanding in their own words (Keuchler and Simpkin, 2010)

- Give faculty greater insight into student thinking compared to multiple choice assessments (Birenbaum and Tatsuoka, 1987)

- Students treat CR and multiple-choice assessments as different cognitive tasks and prepare for them differently (Stanger-Hall, 2012)


AACR Objectives

- Evaluate students’ understanding of scientific concepts
- Create models of student thinking
- Use linguistic and statistical analysis to analyze students’ writing
Question Development Cycle

- Predictive Model
- Confirmatory Analysis
- Human Coding
- Rubric Development
- Question Revision
- Data Collection
- Lexical Resource Development
- New Question Design
- Exploratory Analysis
Data Collection: Study Population

- 3 sections of Introductory Biology Cell and Molecular Course for Majors
- 4 instructors

<table>
<thead>
<tr>
<th></th>
<th>Section 1</th>
<th>Section 2</th>
<th>Section 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enrollment</td>
<td>309</td>
<td>466</td>
<td>302</td>
</tr>
<tr>
<td>% Female</td>
<td>46</td>
<td>58</td>
<td>49</td>
</tr>
<tr>
<td>% First and second years</td>
<td>69</td>
<td>69</td>
<td>67</td>
</tr>
<tr>
<td>Cum GPA at start of term</td>
<td>2.48</td>
<td>2.69</td>
<td>2.52</td>
</tr>
</tbody>
</table>
Scope of Analysis

• 15 questions
  • Genetics, Thermodynamics, Acid-Base Chemistry, Metabolism

• Responses collected
  Pre    Post    Total
  8,290  4,387  12,677
Timeline:
Feedback Report and Just-in-Time Teaching

- Previous week
  - Administer Questions Online

- Monday morning
  - Download Responses, Analyze, Generate Report

- Monday ~5pm
  - Report sent to Instructor

- Tuesday morning
  - Instructors Use Report to Modify Lesson Plan
  - Just-in-Time Instruction

Previous week

Monday morning

Monday ~5pm

Tuesday morning
Question

• Using your knowledge of genetics, explain how human brain cells and heart cells are different.

derived from Genetics Concept Assessment (Smith et al., 2008)
Lexical Resource Development: IBM SPSS Text Analysis

Each response is classified into 0 or more categories.
- Connects text and cluster analyses
- Generates output for feedback report
- Allows rapid analysis of new data sets
## Feedback Report

<table>
<thead>
<tr>
<th>Cluster Description</th>
<th>Cluster 1 44%</th>
<th>Cluster 2 28%</th>
<th>Cluster 3 27%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cluster description</td>
<td>Gene expression</td>
<td>Cell function and physiology</td>
<td>Different DNA</td>
</tr>
<tr>
<td>Sample Response</td>
<td>Each cell has different genes turned on and off which expresses different proteins.</td>
<td>The function of the cells within the heart cells and eye cells is what makes them different.</td>
<td>Cells have different jobs Heart cells and eye cells are different because their DNA is different the DNA is coded into different RNA strands which code for proteins that are very different…</td>
</tr>
</tbody>
</table>
# Distribution of Categories by Cluster

<table>
<thead>
<tr>
<th>Lexical category</th>
<th>Cluster 1</th>
<th>Cluster 2</th>
<th>Cluster 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gene expression</td>
<td>98%</td>
<td>6%</td>
<td>11%</td>
</tr>
<tr>
<td>Genes</td>
<td>79%</td>
<td>2%</td>
<td>2%</td>
</tr>
<tr>
<td>DNA</td>
<td>15%</td>
<td>14%</td>
<td>88%</td>
</tr>
<tr>
<td>Physiology</td>
<td>11%</td>
<td>66%</td>
<td>17%</td>
</tr>
<tr>
<td>Protein</td>
<td>60%</td>
<td>10%</td>
<td>46%</td>
</tr>
<tr>
<td>Different</td>
<td>92%</td>
<td>75%</td>
<td>94%</td>
</tr>
</tbody>
</table>

>70% of responses were assigned to the category

30-70% of responses were assigned to the category

<30% of responses were assigned to the category
Gene expression

Cell function and physiology

Different DNA

Graphs showing relationships between DNA, Gene expression, Protein, Physiology, Gene, and Different. Node size indicates the proportion of responses in each category:
- Solid line: Share 75 – 100% of responses
- Dashed line: Share 50 – 74% of responses
- Dotted line: Share 25 – 49% of responses
### Comparison of Student Responses

**PRE/POST Instruction**

<table>
<thead>
<tr>
<th>Initial distribution</th>
<th>Final distribution</th>
<th>Post-instruction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-instruction</td>
<td></td>
<td></td>
</tr>
<tr>
<td>35%</td>
<td>Cluster</td>
<td>Gene expression</td>
</tr>
<tr>
<td></td>
<td></td>
<td>70%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>13%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>17%</td>
</tr>
<tr>
<td>37%</td>
<td>Cell function and physiology</td>
<td>23%</td>
</tr>
<tr>
<td>28%</td>
<td>Different DNA</td>
<td>41%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>18%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>42%</td>
</tr>
</tbody>
</table>
Instructors response to CR questions and feedback reports

• Created clicker questions and led discussions based on results from feedback report

• Reported that written assessments were particularly important for gaining insight as to why students have struggled continuously with certain concepts

• Proposed future in-class activities to improve student writing skills
Improving Automated Analysis for JiTT

- Encourage student participation by giving credit for homework assignments
- Allow more time between each assignment and the next class for preparing instructional activities
- Professional development for faculty to help them address concepts that students find challenging
Future Directions

- Faculty Learning Communities
  - Local: groups of faculty within a department or teaching the same course
  - Virtual: faculty across institutions
- Use the same assessment & share instructional materials
Future Directions

- Web Portal
  - support rapid assessment and feedback
Acknowledgements

- Automated Analysis of Constructed Response Research Group (AACR)
  - Michigan State University
  - University of Colorado – Boulder
  - University of Maine
  - SUNY – Stony Brook
  - University of Georgia
  - University of California – LA
  - Western Michigan University

This work was supported by the National Science Foundation (Grant DUE 1022653). Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.