Insights into Student Thinking in STEM: Lessons Learned from Lexical Analysis of Student Writing

Automated Analysis of Constructed Response Research Group
aacr.crcstl.msu.edu

Michigan State University

Mark Urban-Lurain
Rosa A. Moscarella
Kevin C. Haudek
Emma Giese
Duncan F. Sibley
John E. Merrill

NARST Paper S12.13.4 March 24, 2010
Acknowledgements

- National Science Foundation DUE-0736952, DUE-0243126 and EHR-0314866
- Carnegie Corporation Grant #B7458
- The Vice Provost for Libraries, Computing and Technology, Michigan State University

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 funding agencies.
Overview

- Introduction
- Methods
- Example Results
- Lessons Learned
Assessment Design Constraints

- Conceptual barriers impair students’ understanding of complex processes in science
  - May be identified by students’ use of language
  - *Constructed Response* questions can provide insight into learning obstacles

- Large courses prohibit using constructed response questions
  - How can we still achieve our assessment goal?
Assessment Types

Closed

- Multiple Choice
- Selection/Identification
- Reordering/Rearrangement
- Completion
- Construction

Constructed

- Presentation
Objectives

- Use linguistic analysis software to analyze students’ written responses
  - Develop necessary libraries and resources
- Evaluate students’ understanding of various scientific concepts using these tools
  - Predict expert ratings
Computerized Lexical Analysis Approaches

- Linguistic Feature-based Methods
- Vector Space Methods
- Linguistic Structure Analyses
Our Approach: Linguistic Feature-Based

- Item Construction
- Disciplinary Term Extraction
- Disciplinary Construct Identification
- Expert Scoring
- Statistical Modeling

- Student Responses
SPSS Text Analysis for Surveys

The image shows a screenshot of SPSS Text Analytics for Surveys, which is used for categorizing responses based on themes or categories. The interface includes a window with a table of responses and categories, such as 'energy', 'fat', 'calories', 'biological process', 'cell', 'carbon dioxide', 'converted', and 'metabolism'. The text analysis highlights themes like metabolism, fat stores, and energy expenditure in the context of nutrition and physiology.
Examples
Tracing Carbon in Cellular Respiration

<table>
<thead>
<tr>
<th>Cells in an active muscle release CO2. How did the carbon get into the CO2?</th>
<th>What substance was the carbon in?</th>
<th>How did it get there?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Start here</td>
<td><strong>Carbon Dioxide</strong></td>
<td></td>
</tr>
<tr>
<td>Before that...</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Before that...</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Before that...</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- Categorize compounds and processes to find patterns in student thinking

Predicting Expert Ratings

- Acid/base chemistry of biological functional groups
  - Explanation of a strong acid and a weak acid.
- Three category scoring rubric by two experts
- Use categories identified by lexical analysis to predict expert ratings
- Classified 83.8% of all cases correctly
  - Inter-rater reliability between experts and computer predictions 0.882

Weight Loss: Multiple Choice

- DQC question
  - BS 111, fall, 2006
  - N = 459

You have a friend who lost 15 pounds of fat on a diet. Where did the mass go?

44.2% A) The mass was released as CO2 and H2O.
23.3% B) The mass was converted to energy which was used up.
21.1% C) The mass was converted to ATP molecules.
8.7% D) The mass was broken down to amino acids and eliminated from the body.
2.6% E) The mass was converted to urine and feces and eliminated from the body.
You have a friend who lost 15 pounds of fat on a diet. Where did the mass go?
### Distribution of Categories

<table>
<thead>
<tr>
<th>Category</th>
<th>Selection %</th>
<th>Respondents</th>
<th>Total %</th>
</tr>
</thead>
<tbody>
<tr>
<td>energy</td>
<td>69.0</td>
<td>218</td>
<td>69.0</td>
</tr>
<tr>
<td>fat</td>
<td>59.5</td>
<td>188</td>
<td>59.5</td>
</tr>
<tr>
<td>calories</td>
<td>54.1</td>
<td>171</td>
<td>54.1</td>
</tr>
<tr>
<td>metabolism</td>
<td>53.2</td>
<td>168</td>
<td>53.2</td>
</tr>
<tr>
<td>carbon dioxide</td>
<td>48.7</td>
<td>154</td>
<td>48.7</td>
</tr>
<tr>
<td>water</td>
<td>46.5</td>
<td>147</td>
<td>46.5</td>
</tr>
<tr>
<td>converted</td>
<td>24.4</td>
<td>77</td>
<td>24.4</td>
</tr>
<tr>
<td>biological process</td>
<td>20.9</td>
<td>66</td>
<td>20.9</td>
</tr>
<tr>
<td>atp</td>
<td>19.3</td>
<td>61</td>
<td>19.3</td>
</tr>
<tr>
<td>sweat</td>
<td>17.1</td>
<td>54</td>
<td>17.1</td>
</tr>
<tr>
<td>heat</td>
<td>16.8</td>
<td>53</td>
<td>16.8</td>
</tr>
<tr>
<td>burned</td>
<td>16.1</td>
<td>51</td>
<td>16.1</td>
</tr>
<tr>
<td>urine &amp; feces</td>
<td>15.8</td>
<td>50</td>
<td>15.8</td>
</tr>
<tr>
<td>cell</td>
<td>12.7</td>
<td>40</td>
<td>12.7</td>
</tr>
<tr>
<td>waste</td>
<td>12.7</td>
<td>40</td>
<td>12.7</td>
</tr>
<tr>
<td>fuel</td>
<td>9.2</td>
<td>29</td>
<td>9.2</td>
</tr>
<tr>
<td>As energy</td>
<td>7.9</td>
<td>25</td>
<td>7.9</td>
</tr>
</tbody>
</table>

18 Categories
315 (99%) Responses Categorized
0 0
The cartoon above represents change that has occurred in a population of animals and a population of plants over thousands of years (time is read from left to right). Use your current understanding of evolution by natural selection to explain how the changes came about.

## Constructed Response Assessment Framework

<table>
<thead>
<tr>
<th>Question Structure</th>
<th>Answer Length/Complexity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Short/Single Concept</td>
</tr>
<tr>
<td><strong>More Directed/Structured</strong></td>
<td>Box and Arrow, SBF Models</td>
</tr>
<tr>
<td><strong>Less Directed/Structured</strong></td>
<td>DQC, Analogies</td>
</tr>
</tbody>
</table>
Lessons Learned
Question Structure

- Simple questions can produce complex answers
  - “Explain the difference between a weak and strong acid.”

- Decompose questions
  - Give an example of a strong acid
  - Explain strong acids
  - Give an example of a weak acid
  - Explain weak acids
Question Wording

- Words in question will likely be repeated in responses
- Select question words to distinguish correct from incorrect responses
  - Should not include key words
  - Have target scientific terminology to identify in responses
Response Length

- Single word (or molecular formula) to several sentences
- Long enough for extracted terms for accurate categorization.
  - Too long, too many unrelated terms
  - Too short, may only in one or two categories
    - Hard to see connections students make between concepts
- Directed questions produce responses consisting of one to two sentences
  - More easily categorized
Lexical Categories

- Categorization should be fine grained
  - Can collapse for further analysis
  - Statistical prediction requires fewer
    - Difficult to predict which ones \textit{a priori}
- Some answers cannot be categorized
  - Usually lack of understanding
  - Brain dump, word salad
- Expert classification rubric
  - Holistic vs. structured
  - Granularity
Building and Sharing Custom Libraries

- Multiple libraries
  - Metabolism, genetics, evolution, geology, verbs
  - Verbs not extracted by default
  - Include inflections - term extraction more efficient
    - Evolve: Evolves, evolving, and evolved

- Make changes in Local Library
  - Librarian regularly merges, then publishes and distributes updated libraries
Summary

- Multiple choice questions don’t tell whole story
- Lexical analysis provides a whole-class picture of term / concept usage
- Statistical analysis can help identify categories of importance
Questions

Automated Analysis of Constructed Response Research Group web site
http://aacr.crcstl.msu.edu

Mark Urban-Lurain
Division of Science and Mathematics Education
Michigan State University
urban@msu.edu