Examining students' ability to apply thermodynamics concepts in biology: computerized lexical analysis reveals heterogeneous ideas

Luanna Prevost[1], Kevin Haudek[1], John Merrill[2], Mark Urban-Lurain[1]
[1]Center for Engineering Education Research, [2]Biological Sciences Program, Michigan State University, East Lansing, MI 48823

Introduction

- Thermodynamics concepts are foundational to the STEM fields, but challenging to students(Streveler et al. 2011)
- Constructed Response (written) assessments can provide insight into student ideas (Birenbaum and Tatsuoka, 1993)
- We use a linguistic feature-based approach for automated analysis of student writing
- We evaluate introductory biology students' understanding of thermodynamic at Bloom's *comprehension* and *application* levels

Question 1 :Comprehension

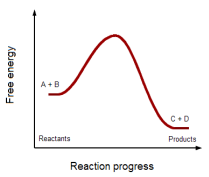
Study the graph below that illustrates the chemical reaction A +B → C + D. Does this graph represent an exergonic or endergonic reaction?

A.Endergonic

B

C

H



84.6%

Explanations from correct MC responses were scored by experts and used for text analysis

Question 1 :Human Scoring Rubric

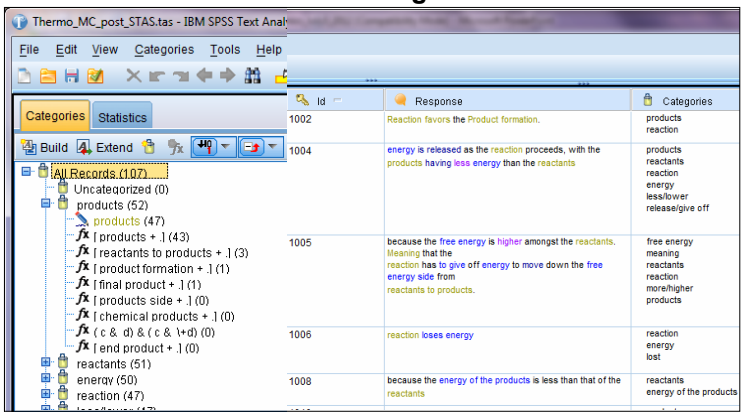
Level	% (n=168)	Rubric	Example
Correct	49	Totally correct explanation	energy is released as the reaction proceeds, with the products having less energy than the reactants
Incorrect	51	incorrect or incomplete explanation	The energy is transferred from the reactants to the products

Two expert interrater reliability=0.88 intraclass correlation

References:
Birenbaum, M., and K. K. Tatsuoka. 1987. *Open-Ended Versus Multiple-Choice Response Formats—It Does Make a Difference for Diagnostic Purposes*. Applied Psychological Measurement 11:385 –395.
Chi, M. T. H., P. J. Feltovich, and R. Glaser. 1981. *Categorization and representation of physics problems by experts and novices*. Cognitive science 5:121–152.
Deane, P. 2006. Strategies *for evidence identification through linguistic assessment of textual responses*. Pages 313–371 in D. M. Williamson, I. I. Bejar, and R. J. Mislevy, editors. Automated scoring of complex tasks in computer based testing.
Streveler, R. A., R. L. Miller, A. I. Santiago-Román, M. A. Nelson, M. R. Geist, and B. M. Olds. 2011. *Rigorous Methodology for Concept Inventory Development: Using the "Assessment Triangle" to Develop and Test the Thermal and Transport Science Concept Inventory (TTCI)*. International Journal of Engineering Education 27:968.

Acknowledgements
This material is based upon work supported by the National Science Foundation (DUE-ASA 0243126, DUE 0736952 and 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 supporting agencies.

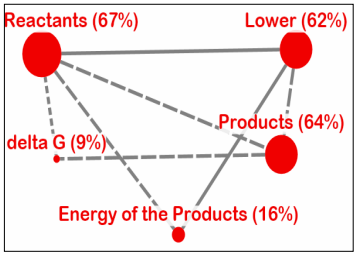
Question 1: Text Analysis Extraction and Category Building



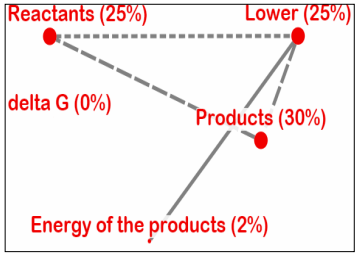
IBM SPSS Text Analytics for Surveys software showing the terms extracted (colored words), categories (upper left panel), student responses (right panel). Each response is placed into one or more categories (rightmost column). A total of 15 lexical categories were created.

Question 1: Correct responses have more correct ideas which are more connected

Correct



Incorrect



Question 1 Key: Node size is relative to percent of total responses in that category
Solid: ——— Share 75 – 100 % of responses
Dashed: - - - - Share 50 - 74% of responses
Dotted: ------- Share 25 – 49% of responses

Discriminant Analysis Results: Categories predictive of expert scoring

Category	Standardized Coefficient
Delta G	0.525
Energy of the products	0.502
Products	0.492
Lower	0.469
Reactants	0.360

Computer models correct predict 70% of expert scoring

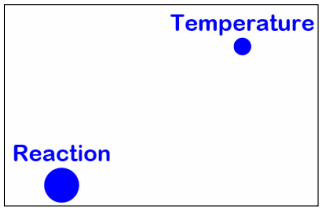
Question 2: Application

A carbohydrate is composed of a string of covalently linked monosaccharides. Breaking those bonds between the monosaccharides is a chemically spontaneous reaction (ΔG for this reaction is -3.7 kcal/mol). However, this reaction occurs very slowly at room temperature. Why do you think this is so?

Question 2: Web diagrams show relationship among common ideas (categories) in each cluster

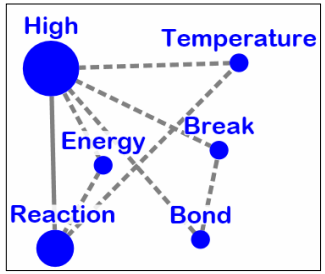
- Categories were extracted using text analysis
- K-means Cluster Analysis groups responses with similar characteristics from text analysis categories

Cluster #1



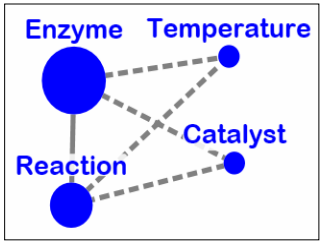
- 72% were assigned to cluster 1 with no apparent pattern

Cluster #2



- 18% expressed ideas about bond stability and the need for higher temperatures to increase the reaction rate

Cluster #3



- 10% described that a catalyst would help increase the rate
- students referenced biological knowledge (a "surface" feature) rather than thermodynamic knowledge in addressing thermodynamics concepts ("deep" feature") (Chi et al 1981)

Who we are: The Automated Analysis of Constructed Response (AACR) research group consists of researchers from multiple universities with backgrounds in various STEM disciplines, linguistics, and educational research. Other institutions include: Ohio State University; University of Colorado – Boulder, University of Georgia, University of Maine.
Visit us at **http://aacr.crcstl.msu.edu/**