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

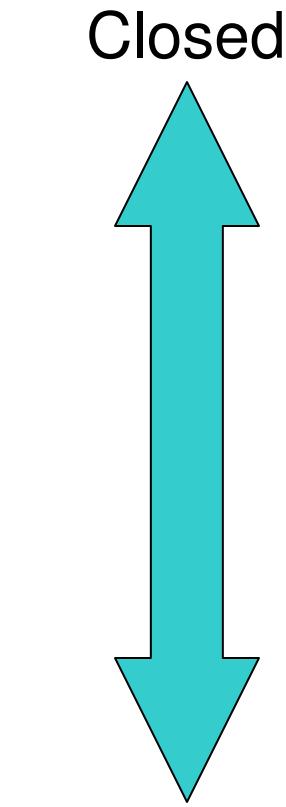
Overview

- Introduction
- Methods
- Example Results
- Lessons Learned

Assessment Design Constraints

- Conceptual barriers impair students' understanding 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



Multiple Choice

Selection/Identification

Reordering/Rearrangement

Completion

Construction

Presentation

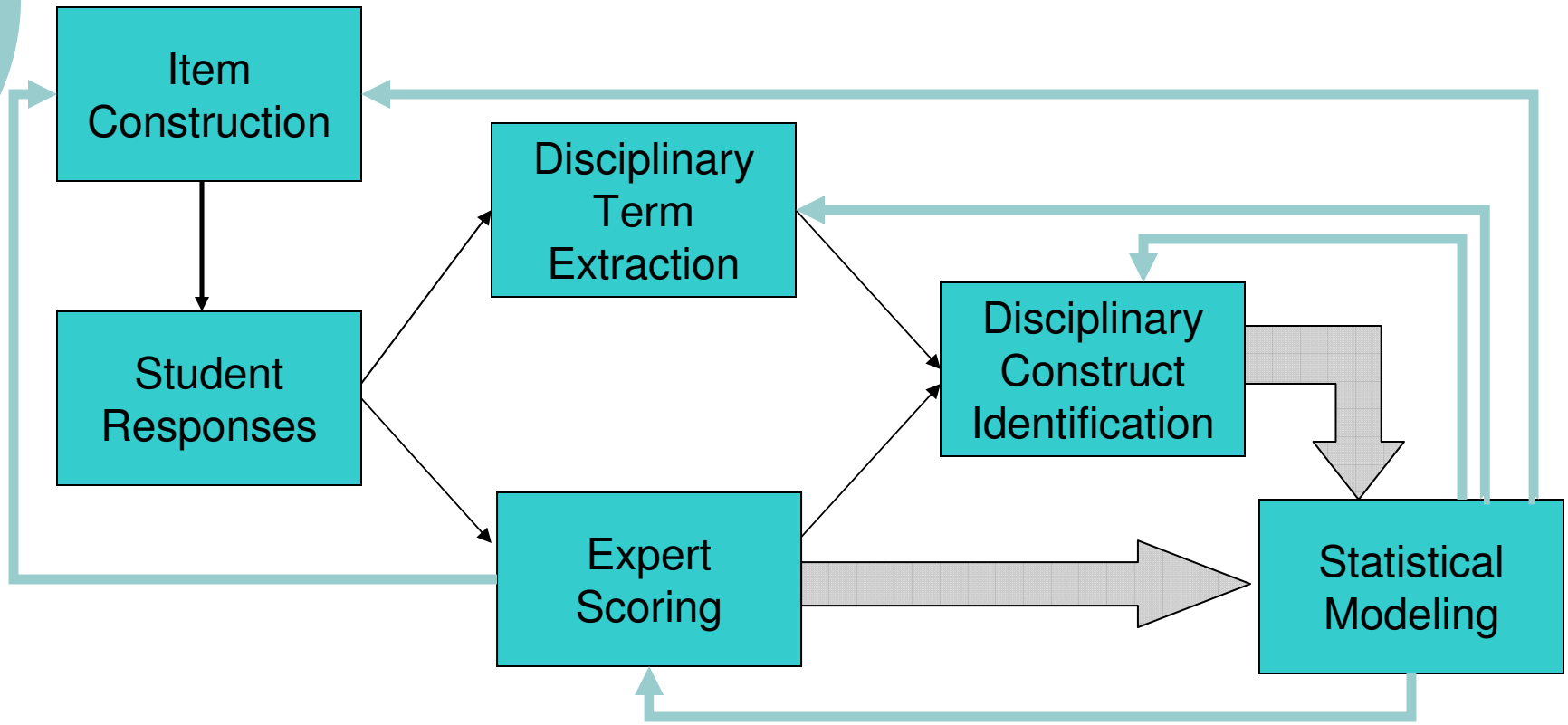
Objectives

- 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

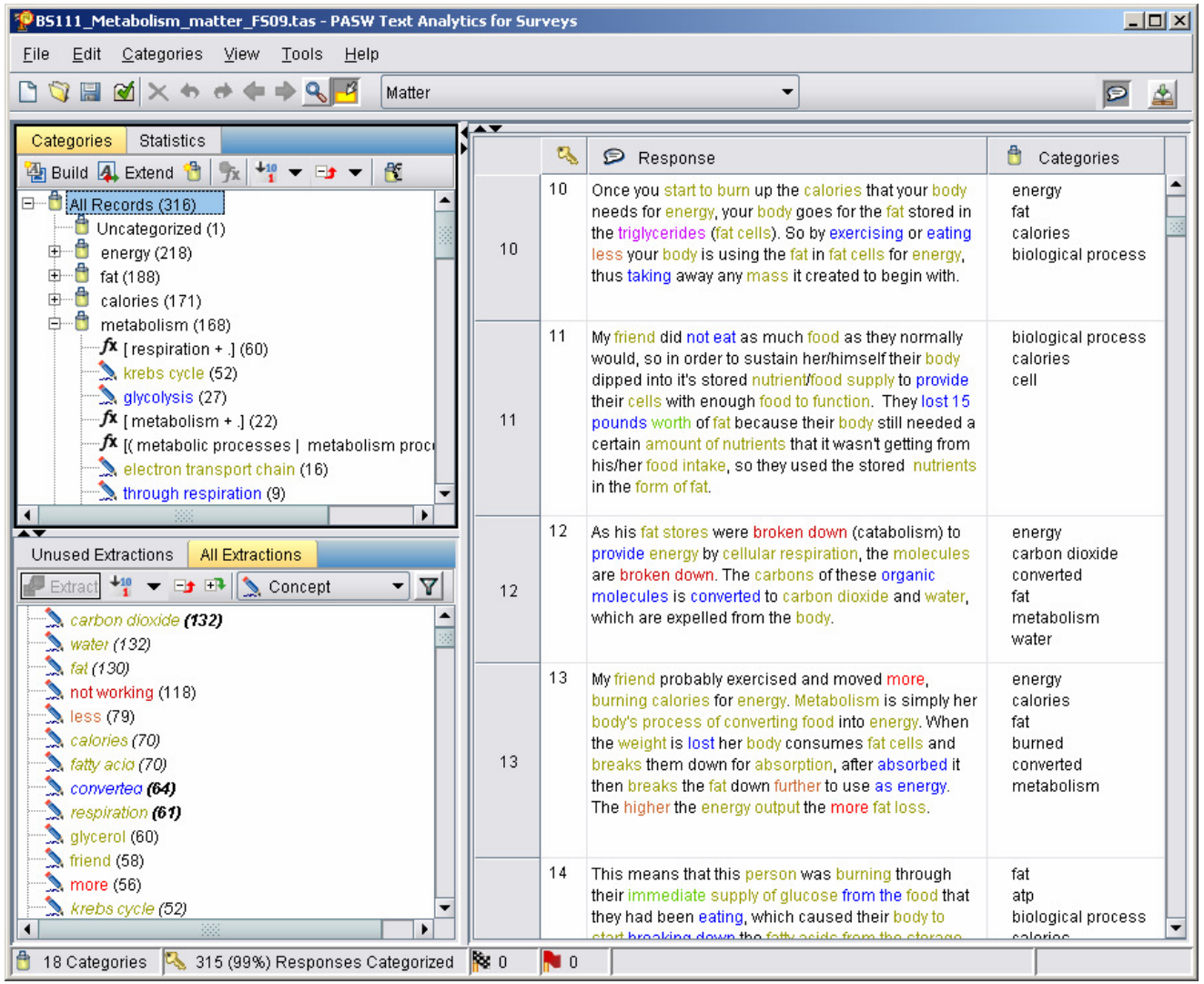
Computerized Lexical Analysis Approaches

- Linguistic Feature-based Methods
- Vector Space Methods
- Linguistic Structure Analyses

Our Approach: Linguistic Feature-Based



SPSS Text Analysis for Surveys



Examples

Tracing Carbon in Cellular Respiration

Cells in an active muscle release CO_2. How did the carbon get into the CO_2?	**What <u>substance</u> was the carbon in?**	**<u>How</u> did it get there?**
Start here	*Carbon Dioxide*	
Before that…		
Before that…		
Before that…		

- Categorize compounds and processes to find patterns in student thinking

Moscarella, R. A., Urban-Lurain, M., Merritt, B., Long, T., Richmond, G., Merrill, J., et al. (2008, March 30 - April 2). *Understanding undergraduate students' conceptions in science: Using lexical analysis software to analyze students' constructed responses in biology.* Paper presented at the NARST 2008 Annual International Conference, Baltimore, MD.

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

Haudek, K., Moscarella, R. A., Urban-Lurain, M., Merrill, J., Sweeder, R., & Richmond, G. (2009, April 17-21). *Using lexical analysis software to understand student knowledge transfer between chemistry and biology.* Paper presented at the National Association of Research in Science Teaching Annual Conference, Garden Grove, CA.

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 CO_2 and H_2O.

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.

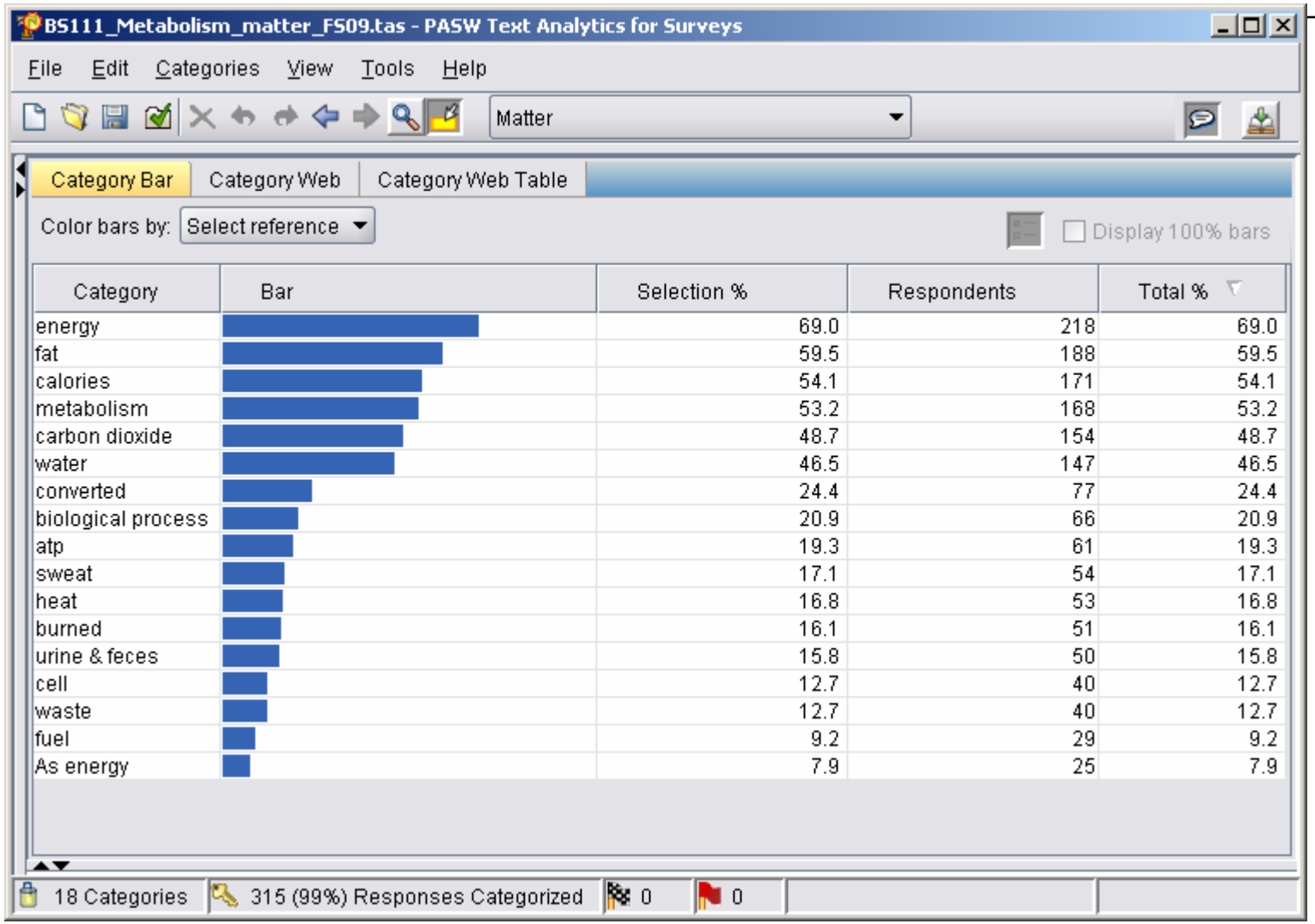
2.6% E) The mass was converted to urine and feces and eliminated from the body.

Weight Loss: Constructed Response

- BS 111, fall, 2009
  - N = 316

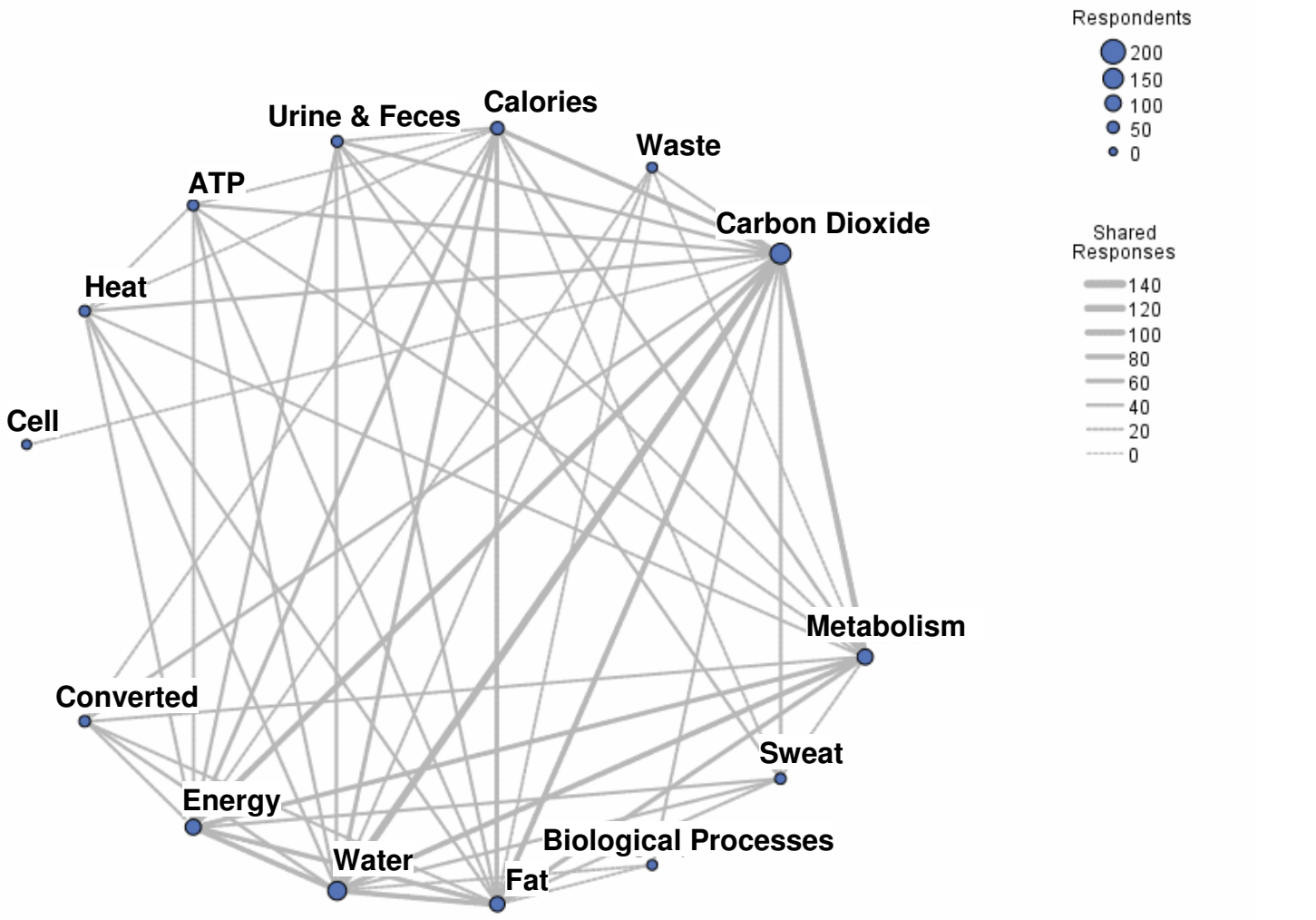
You have a friend who lost 15 pounds of fat on a diet. Where did the mass go?

Distribution of Categories

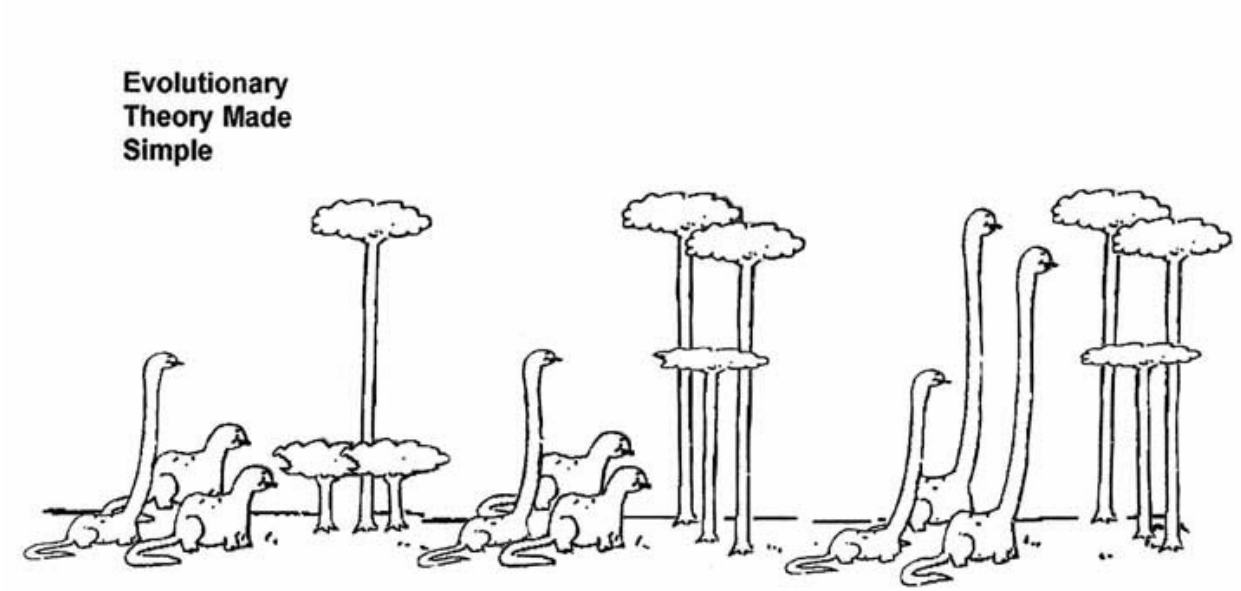


Category	Bar	Selection %	Respondents	Total %
energy		69.0	218	69.0
fat		59.5	188	59.5
calories		54.1	171	54.1
metabolism		53.2	168	53.2
carbon dioxide		48.7	154	48.7
water		46.5	147	46.5
converted		24.4	77	24.4
biological process		20.9	66	20.9
atp		19.3	61	19.3
sweat		17.1	54	17.1
heat		16.8	53	16.8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

Relationships Among Categories



The Dino Problem



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.

Bray Speth, E., Long, T., Pennock, R., & Ebert-May, D. (2009). Using Avida-ED for teaching and learning about evolution in undergraduate introductory biology courses. *Evolution: Education and Outreach, 2*(3), 415-428.

Constructed Response Assessment Framework

		Answer Length/ Complexity	
		Short/ Single Concept	*Long/ Explanatory*
Question Structure	*More Directed/ Structured*	Box and Arrow, SBF Models	MC ➔ Explain your answer
	Less Directed/ Structured	DQC, Analogies	Dino Problem

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 *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