Lexical and qualitative analysis of students' written responses and interviews explaining chemical functional group behavior in cellular biology

Kevin C. Haudek
Luanna Prevost
John Merrill
Mark Urban-Lurain

Center for Engineering Education Research

MICHIGAN STATE UNIVERSITY

ceer Center for Engineering Education Research @ Michigan State University

Acknowledgements

- AACR research group
- National Science Foundation (DUE 1022653)
- Rosa Moscarella

www.aacr.crcstl.msu.edu

Introduction

- Students often have heterogeneous ideas
- Constructed response questions afford a unique look into student thinking
- But their use is often constrained in large enrollment courses by evaluation limits

Automated Analysis of Constructed Response (AACR) Research Group

- We have begun to use lexical analysis to evaluate students' constructed responses in various STEM disciplines
- These results also give insight into student thinking and allow multiple ideas to be detected
- Lexical analysis results can be coupled to statistical functions for classification

Chemistry in biology

- Introductory biology course – *Cells and Molecules*
- *Consider two small chemical compounds that are identical except one contains a hydroxyl group (-OH) and one contains an amino group (-NH_2); if you treated two sets of cells, one set with each compound and measured the cytoplasmic pH, which compound would have the greater effect on the pH of the cytoplasm?*

  A. Amino B. Hydroxyl C. Both D. Neither

- *Explain your answer.*

Haudek et al. (In Press) What are they thinking? Automated analysis of student writing about acid/base chemistry in introductory biology. *CBE-Life Sci Educ.*

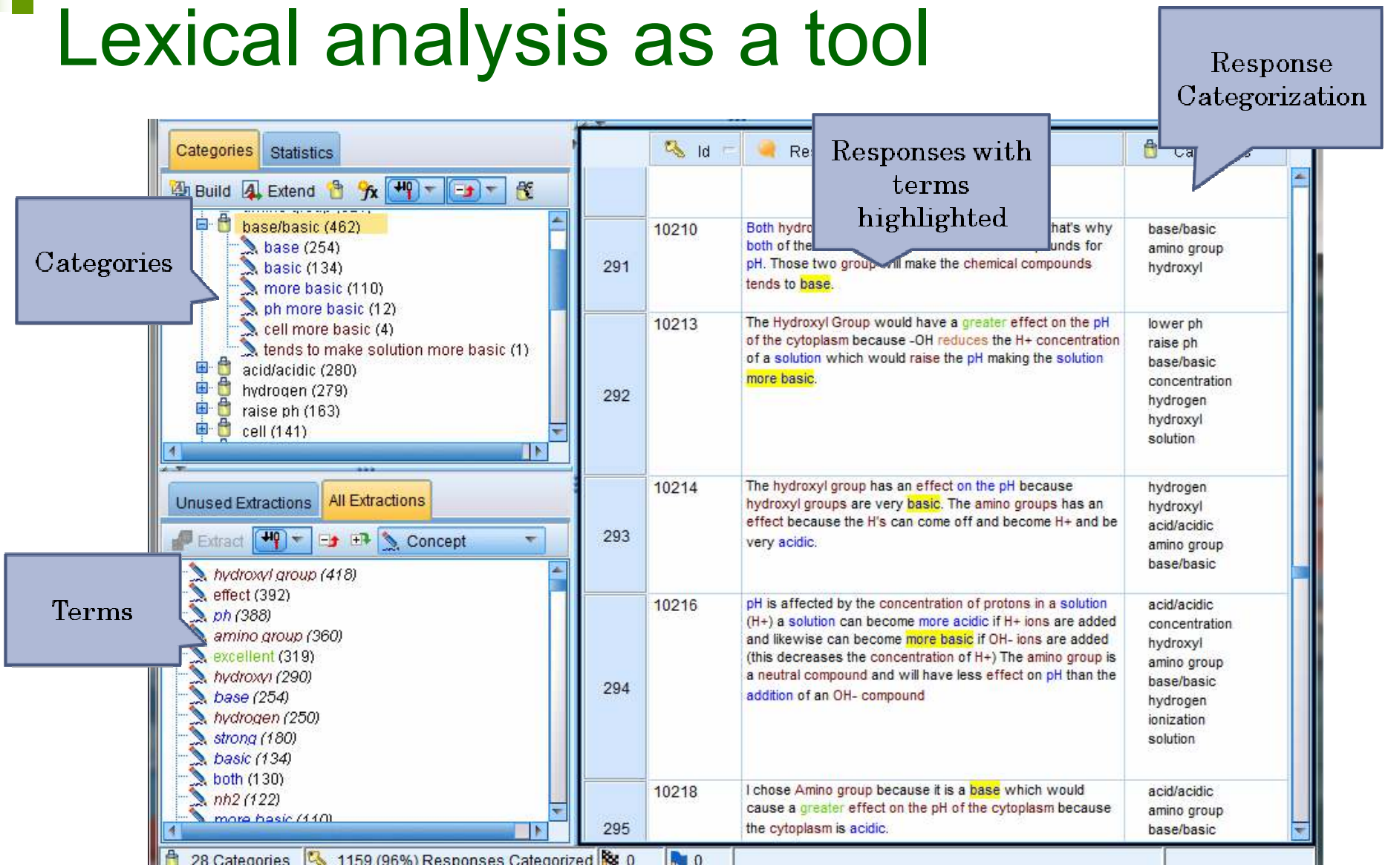
Research questions

- What kind of mental models do students carry that lead to difficulties?
- How well do students' short responses reveal their thinking?

Methods

- Introductory biology course
  - generally STEM majors and 1st/2nd year in college
- During course: Online homework assignment
  - Multiple attempts allowed; credit awarded for completion
  - 179 students completed HW question (460 enrolled) with 195 unique submissions
- After course: Face-to-face interviews
  - Students who selected the correct MC answer (58) were contacted for an interview; 13 students responded
  - Interviews transcribed
- Both explanations subjected to qualitative and lexical analysis
  - IBM SPSS Text Analytics for Surveys, v.4.0

Lexical analysis as a tool



Students' homework explanations

Category	Number of responses
hydroxyl	129
amino	103
base	84
acid	51
hydrogen	39
cell	30
compound	27
raise pH	26
strong base	23
ionization	22
solution	20
accept hydrogen	11
lower pH	10

Category	Number of responses
charge	8
electron	7
hydroxide	7
guess	6
weak base	6
concentration	5
nitrogen	4
strong acid	3
hydrogen bond	2
weak acid	2
alcohol	1

n=195

Interview structure

- Think-aloud protocol
  - Asked specific HW question again
- Specific probing
  - Content
  - Strategy
  - Question stem

Comparing interviews and homework

Multiple choice selection of first submission

Homework choice

Interview choice	Amino	Hydroxyl	Both	Neither	Total
Amino	**5**	1	0	1	7
Hydroxyl	4	**1**	0	0	5
Both	0	1	**0**	0	1
Neither	0	0	0	**0**	0
Total	9	3	0	1	

Explanation of first submission: Paired t-test

	Word count	Category count
Homework	35.1 +/- 16.1 *	4.7 +/- 1.8
Interview	61.1 +/- 37.7 *	4.9 +/- 2.4

* $p < 0.05$

n=13

Lexical categorization of homework and interview explanations

- Students' first HW submission and their verbal response to the think-aloud protocol were subjected to lexical analysis.

Category	HW	Interview		Category	HW	Interview
Accept hydrogen	1	1		Guess	0	5
Acid	3	5		Hydrogen	5	5
Alcohol	1	0		Hydroxyl	12	8
Amino	8	9		Ionization	12	11
Base	5	5		Lower pH	1	1
Cell	2	2		Nitrogen	0	1
Charge	0	1		Raise pH	2	2
Compound	3	3		Solution	3	0
Concentration	1	0		Strong acid	0	1
Electron	0	1		Strong base	2	2

n = 13

Explanation correlations

- Correlation between lexical categories of HW and interview explanation

Lexical category	Kendall's tau B
Strong base	1.000
Raise pH	1.000
Lower pH	1.000
Accept hydrogen	1.000
Acid	0.693
Ionization	0.677
Hydrogen	0.675
Hydroxyl	0.365
Base	0.350
Amino	0.158
Compounds	0.133
Cell	-0.182

Category in red: $p<0.01$ n=13
Category in blue: $p<0.05$

Content probing: Exposing mental models

- Previously identified mental models of acid-base behavior were frequently uncovered during qualitative analysis:
  - Character Model - specific words or names of groups relates to acidity/basicity
    - #019 Written: *…but i think OH group would have the most affect on pH. The presence of OH groups makes a solution more basic.*
  - Chemical Symbol Model - consider the quantity of H or OH in a chemical formula
    - #015 Written: *Both the OH(-) and the H(+) will affect pH, but because the amino group has two H, it will have twice as strong an effect as hydroxyl.*
- Hydroxide / hydroxyl confusion

Lin, J. W. and M. H. Chiu (2007). "Exploring the characteristics and diverse sources of students' mental models of acids and bases." *International Journal of Science Education* **29**(6): 771-803.

Mental model consistency

- Most students demonstrated reasonable consistency between mental models used in homework and interview

  - Chemical symbol model
  - #022 Written*: I choose the amino group containing more hydrogens. The more hydrogens in that is available.*
  - #022 Interview*: …My guess would be the amino group which would be NH2 has another or an additional hydrogen making it more acidic…*

  - #023 Written: *The amino functional group acts as a base and therefore would affect the pH of a cell more than an OH functional group the would be an alcohol.*
  - #023 Interview: …*Cytoplasmic pH…let's see…the compound with the amino group just because that would make it basic. Because I know it's a weak base, because I learned that in chemistry…*

Content probing : Justification

- A few students gave a different answer in the interview than their HW.
- Students tended to justify HW answer as more correct.
  - #021 Interview: *Probably my first since it was fresh in my memory.*
  - #013 Interview: *I've just memorized the strong bases as opposed to the weak ones.*

    INTERVIEWER: *Okay. Do you remember any examples?*

    #013 Interview: *I know which is a strong base, for sure. NH3 is a weak base.*

Content probing : Definitions

- Eight students described acids/bases based on a pH range:
  - #015 Interview: *An acid has a pH of less than seven, and a base has a pH of greater than seven.*

- Five students used Arrhenius or Brønsted-Lowry type definitions:
  - #022 Interview: *Typically an acid, depending on its strength, is more or less described as the willingness to give up a proton in the form of an H plus.*

Conclusions

- Students' short written responses align very well with their verbal explanations
  - Both lexical and qualitative analyses support this
- Students use previously reported mental models of acid-base chemistry to explain biological problems
  - Difficulties center around the chemical symbol OH- and number of H+
  - Definitions often rely on surface features (pH range, "lists" of molecules) instead of properties

Questions

- Kevin Haudek
  - haudekke@msu.edu
- Mark Urban-Lurain (PI)
  - urban@msu.edu

www.aacr.crcstl.msu.edu