Comparing Formative Feedback Reports: Human and Automated Text Analysis of Constructed Response Questions in Biology

Michele Weston, Joyce Parker, & Mark Urban-Lurain

Automated Analysis of Constructed Response research group (AACR), Michigan State University

www.msu.edu/~aacr

Introduction

- Constructed response (CR) questions require students to create a written answer from their understanding of the question and the concepts involved in it (Kuechler & Simkin 2010)
- Students treat multiple choice questions and CR questions as different cognitive tasks (Stanger-Hall 2012)
- We have computerized text analysis resources developed for exploratory research in student writing
- We are interested in applying those resources to CR question formative assessment for instructors

Research Questions

- How do instructors analyze answers to CR questions?
- How do the results from automated text analysis techniques compare with the instructors' analyses?

Methods

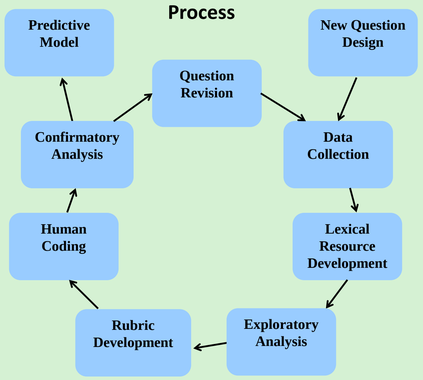
Not all cells in plants (e.g. root cells) contain chlorophyll required for photosynthesis. How do these cells get energy? (Parker et al. 2012)

- Introductory biology course: Cells and Molecules
- Post-instruction on cell metabolism
- Given as a homework on an online course management system
- 360 out of 468 students responded

Instructor Analyses

- Expert scorers experienced with reading student written responses . Had not seen responses to this question before.

Assessment Development Process



Results

Exploratory Analysis

Instructor 1		Instructor 2		Automated Analysis	
Distribution of Responses	Description of Cluster	Distribution of Responses	Description of Cluster	Distribution of Responses	Description of Cluster
15%	1. Accurate description based on the transport of glucose and/or cellular respiration	22%	1. Sugar being transported	13%	1. Sugar being transported through the plant
20%	2. Transport of energy	14%	2. Energy being transferred throughout the plant	23%	2. Energy being used form other parts of the plant
14%	3. Roots drawing something from the soil for energy	22%	3. Energy comes from nutrients from the soil	17%	3. Nutrients from the soil
16%	4. Special processes such as C4 photosynthesis, Calvin Cycle, and Kreb's Cycle	10%	4. Incorrect process such as C4 photosynthesis and heterotrophy	16%	4. Special processes such as electron transport chain and heterotrophy
2%	5. Force-dynamic style explanation	10%	5. Respiration as the process involved	12%	5. Cellular respiration as the process involved
		4%	6. Energy being transferred and nutrients from the soil	16%	6. Incorrect source of energy such as from other organisms

Analytic Scoring Rubric

Instructor 1 Analytic Bins	Instructor 2 Analytic Bins
1. Correct source	1. Transport sugars
2. Incorrect source/unspecified energy	2. Transport energy
3. Incorrect source /nutrients from soil	3. Transport ATP
5. Correct Process	4. Respiration
6. Incorrect Process	5. Dark Reactions
4. Water as Source	6. Soil/nutrients
	7. From surroundings
	8. Transport products of photosynthesis
	9. C4 photosynthesis
	10. Heterotrophy

Confirmatory Analysis/Discriminant Analysis

Rubric	Description	Human Scoring	Correctly Classified
1. Correct Source	Any name for a product of photosynthesis	26%	94.2%
2. Incorrect source/unspecified energy	Energy or ATP being transported	26%	81.7%
3. Incorrect source /nutrients from soil	Energy comes from nutrients from the soil	19%	90.7%
4. Incorrect Source/water	Water without anything else or water with nutrients from the soil	11%	94.1%
5. Correct Process	Respiration or glycolysis	17%	96.8%
6. Incorrect Process	Various incorrect processes	12%	92.6%

Discussion

Time Requirements

Exploratory analysis

- Instructor 1, 11 minutes
- Instructor 2, 45 minutes
- Automated analysis, 15 minutes

Analytic rubric scoring

- Human scoring took approximately 5 hours after calibration
- Automated analysis took less than 5 minutes

Conclusion

How do instructors analyze answers to CR questions?

- Instructors who participated in this project did initial exploratory analyses to look for key ideas and misconceptions
- Then, used those ideas to make analytic scoring rubrics

How do the results from automated text analysis techniques compare with the instructors' analyses?

- 5 of 6 automated clusters overlapped with human clusters
- Discriminant analysis predicted **above 80%** for all the rubrics and **above 90%** for five of them

References

Keuchler, W.L., & Simkin, M.G. (2010). Why is performance on multiple-choice tests and constructed response tests not more closely related? Theory and and empirical test. Decision Sciences Journal of Innovative Education 8:55-73.

Stanger-Hall, K.F. (2012). Multiple-Choice Exams: An Obstacle for Higher-Level Thinking in Introductory Science Classes. *CBE Life Sci Educ, (11)* 3 294-306. doi: 10.1187/cbe.11-11-0100

More Information

A paper for this project can be found at the AACR website.

Contact: Michele Weston at westonmi@msu.edu

Acknowledgements

We thank John Merrill for creating an instructor rubric.

This material is based upon work supported by the National Science Foundation (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 NSF.