Using Computerized Lexical Analysis of Student Writing to Facilitate Just-in-Time Teaching in Large-Enrollment Biology Courses

Luanna B. Prevost, Kevin C. Haudek, Emily Norton Henry, Matthew C. Berry and Mark Urban-Lurain
The Automated Analysis of Constructed Response (AACR) Research Group, Center for Engineering Education Research
www.msu.edu/~aacr

Introduction

Importance of formative & constructed response (CR) assessments
Formative assessments
• Give students feedback on their learning (Pellegrino et al. 2001)
• Allow instructors to modify instruction to improve student learning
Constructive response assessments
• Allow students to represent their understanding in their own words (Keuchler and Simpkin 2010)
• Give faculty greater insight into student thinking compared to multiple choice assessments (Birenbaum and Tatsuoka 1987)
• Students treat CR and multiple-choice assessments as different cognitive tasks and prepare for them differently (Stanger-Hall 2012)

Barriers to using CR assessments
• CR assessments are time consuming to grade, especially in large enrollment courses

Objectives
• Evaluate students’ understanding of scientific concepts
• Facilitate using authentic writing tasks in large STEM courses

Automated Analysis Approach
• Assigns students responses to categories based key words/phrases (Text Analysis)
• Cluster responses based on co-occurrence of categories (K-means Clustering)
• Allows processing of large amounts of data

Automated Analysis using IBM SPSS Modeler

Feedback Report

Within 24 hrs of homework submission deadline instructors received a feedback report with the following information (a–d).

a) Homework question:
Using your knowledge of genetics, explain how human brain cells and heart cells are different.

b) Subset of categories description and frequencies in Clusters 1 and 3

Cluster description
Lexical category
Gene expression
Cell function and physiology
Answers in category
Answers in category
Answers in DNA

Cluster 1
Gene expression
Answers in category
Cell function and physiology
Answers in category
DNA

Cluster 2
Gene expression
6%
14%
46%

Cluster 3
Gene expression
75%
41%

Green= >70% of responses were assigned to the category
Yellow= 30–70% of responses were assigned to the category
Pink =<30% of responses were assigned to the category

Black: no change in correctness of answer
Green: movement towards correct answer
Red: movement away from correct answer

References


Acknowledgments

This material is based upon work supported by the National Science Foundation (DUE 0736952 and DUE 1022553). 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.