

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

Luanna Prevost, Kevin Haudek, Emily Norton Henry,  
Matthew Berry, Mark Urban-Lurain

Automated Analysis of Constructed Response  
(AACR)

Research Group

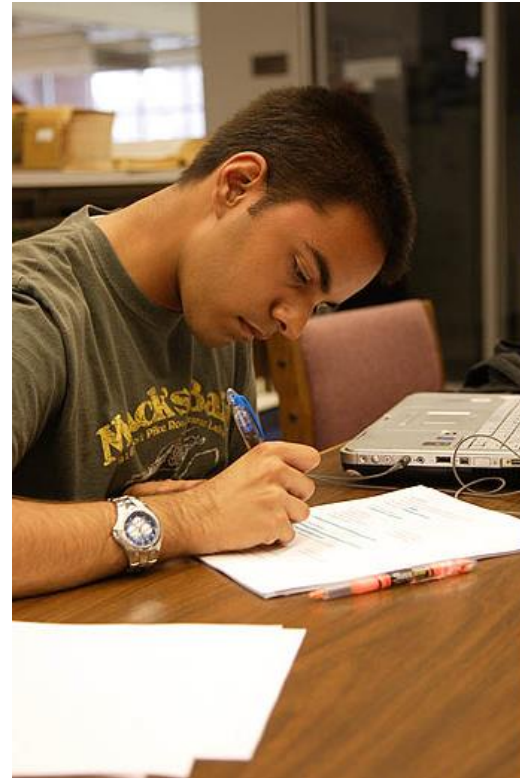
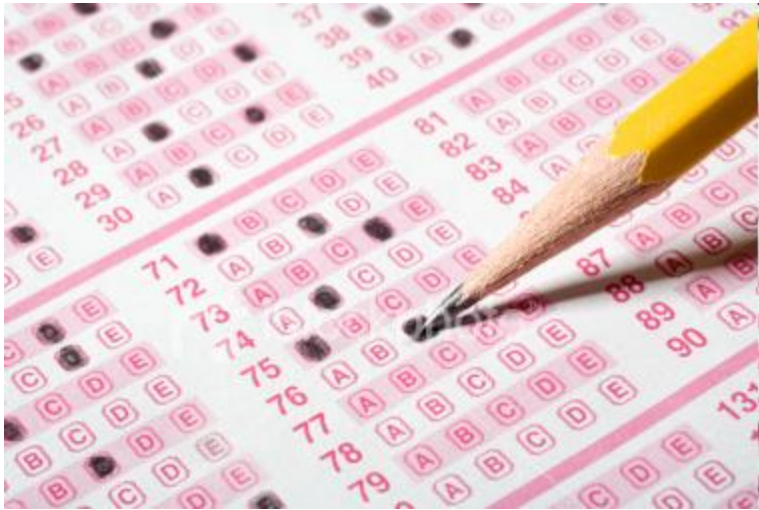
[www.msu.edu/~aacr](http://www.msu.edu/~aacr)



Center for Engineering  
Education Research  
@ Michigan State University

**MICHIGAN STATE**  
UNIVERSITY

# Assessment to Reveal Student Thinking



- Large enrollment courses prohibit the use of constructed response assessments

# Constructed Response (CR) 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)

Kuechler, W. L., & Simkin, M. G. (2010). Why is performance on multiple-choice tests and constructed-response tests not more closely related? Theory and an empirical test. *Decision Sciences Journal of Innovative Education*, 8(1), 55-73.

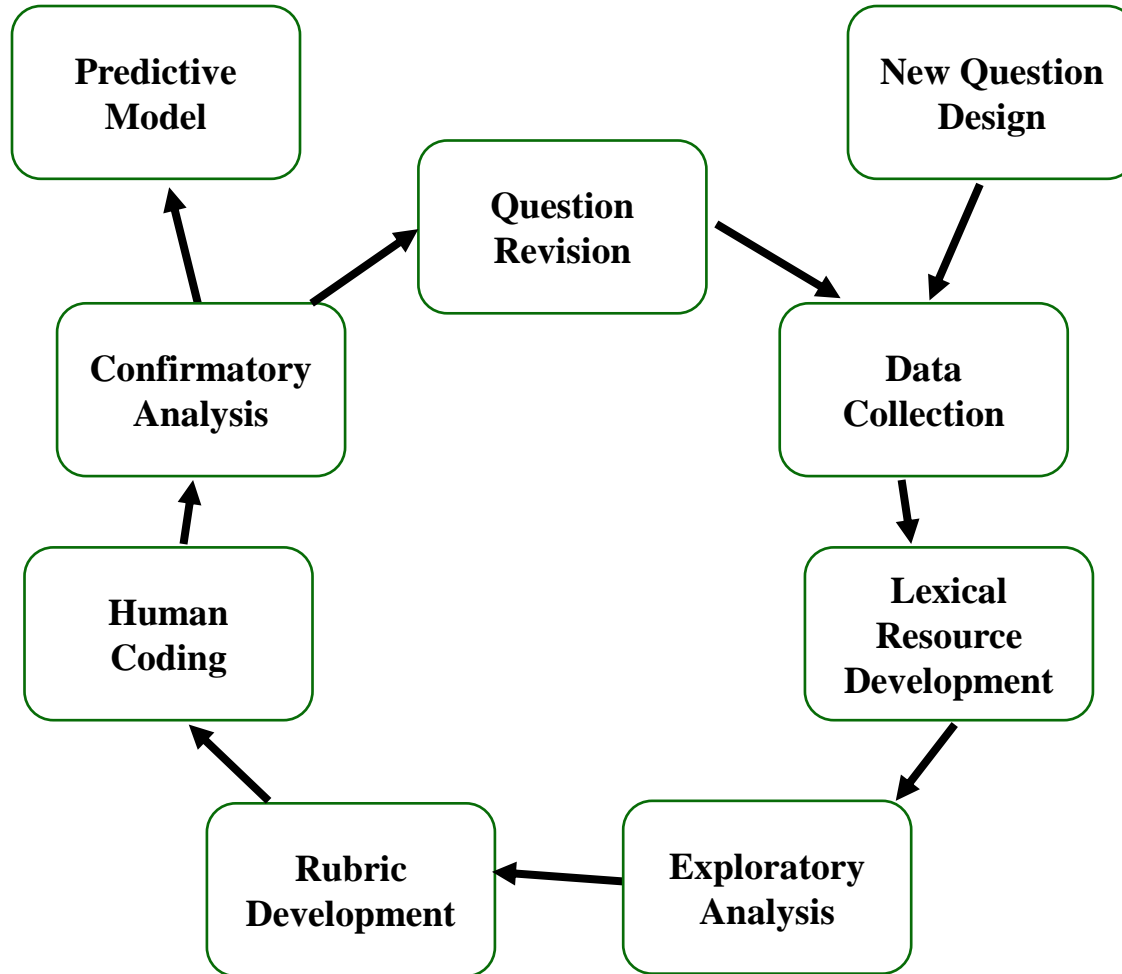
Birenbaum, M., & Tatsuoka, K. K. (1987). Open-ended versus multiple-choice response formats - It does make a difference for diagnostic purposes. *Applied Psychological Measurement*, 11, 329-341.

Stanger-Hall, K. F. (2012). Multiple-choice exams: An obstacle for higher-level thinking in introductory science classes. *CBE-Life Sciences Education*, 11(3), 294-306.

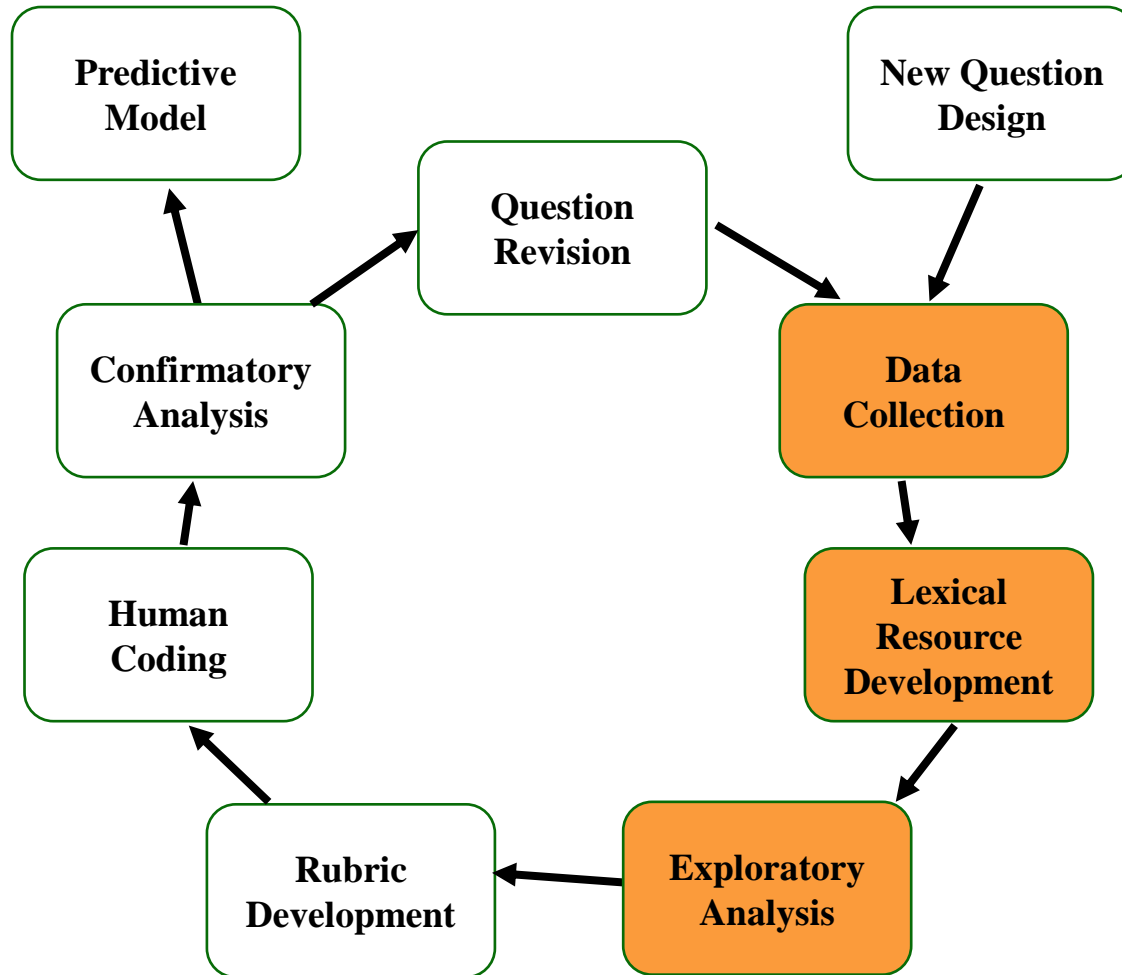
# AACR Objectives

- Evaluate students' understanding of scientific concepts
- Create models of student thinking
- Use linguistic and statistical analysis to analyze students' writing

# Question Development Cycle



# Question Development Cycle



# Data Collection: Study Population

- 3 sections of Introductory Biology Cell and Molecular Course for Majors
- 4 instructors

	Section 1	Section 2	Section 3
Enrollment	309	466	302
% Female	46	58	49
% First and second years	69	69	67
Cum GPA at start of term	2.48	2.69	2.52

# Scope of Analysis

- 15 questions
  - Genetics , Thermodynamics, Acid-Base Chemistry, Metabolism

- Responses collected

Pre

Post

Total

8,290

4,387

12,677



# Timeline: Feedback Report and Just-in-Time Teaching

Previous  
week

Monday  
morning

Monday  
~5pm

Tuesday  
morning

**Administer  
Questions  
Online**

**Download  
Responses ,  
Analyze,  
Generate  
Report**

**Report  
sent to  
Instructor**

**Instructors  
Use Report  
to Modify  
Lesson Plan**

**Just-in-  
Time  
Instruction**

# Question

- Using your knowledge of genetics, explain how human brain cells and heart cells are different.

derived from Genetics Concept Assessment (Smith et al., 2008)

# Lexical Resource Development: IBM SPSS Text Analysis

The screenshot displays the IBM SPSS Text Analysis interface. On the left, a tree view shows 'Lexical categories' with a table of counts. On the right, a table shows 'Student responses' for a question, with each response classified into one or more categories. Green callout boxes highlight these two main components.

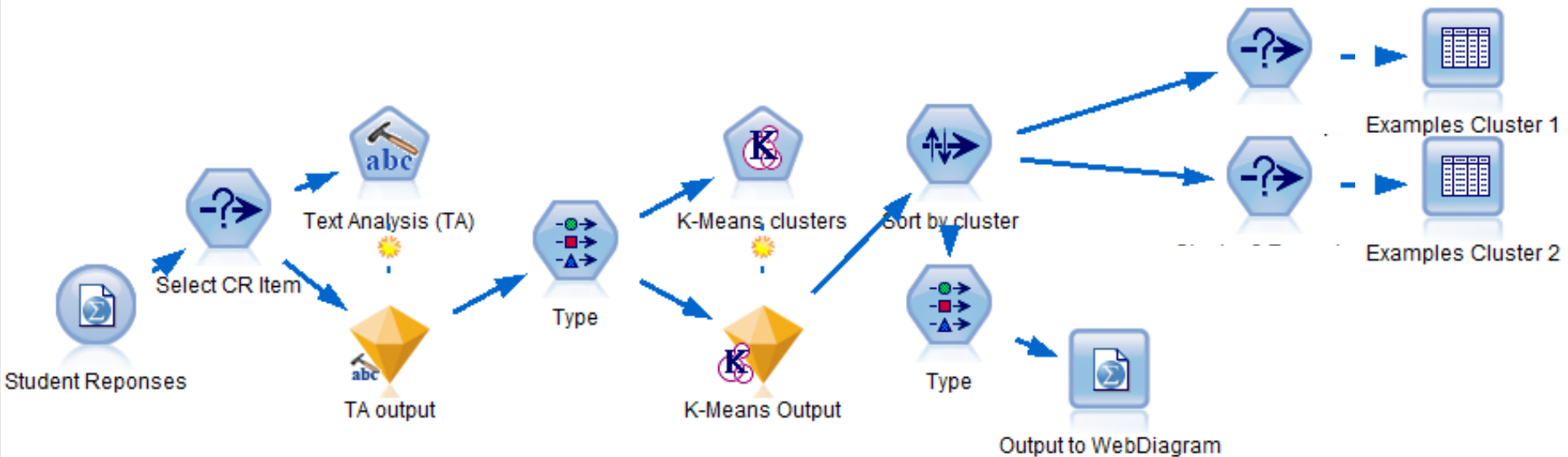
Category	Documents	Docs
All Documents	-	547
Uncategorized	-	10
No concepts extracted	-	0
Alleles	2	2
Chromosomes	1	13
code for	2	59
development	1	3
Different	2	427
differentiated	4	24
Division	8	33
dna	4	167
fx ( dna )		161
dna chain		0
genetic information		5
genetic material		10
Exons and introns	1	1

Response ID	Text	Categories
1	Different parts of the genome are active and elicit different growth factors, the brain cell will not need to withstand as much force as the heart cell so the membrane will not be as thick.	Physiology Different Gene expression Genome
2	The cells would be different because they have different functions. So, in the DNA different proteins and enzymes would be needed for a brain cell when compared to a heart cell.	Different dna function protein
3	They create different sets of proteins by transcribing different genes.	Different protein gene transcribe
4	Brain cells are different from heart cells because they are terminally differentiated and the genes expressed are tightly controlled as to which genes are turned on or off.	Different Gene expression gene differentiated

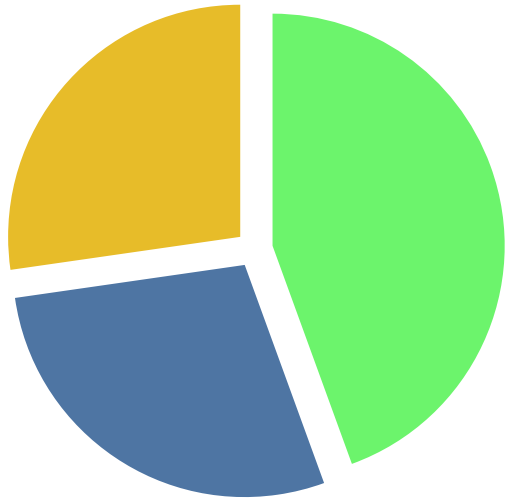
Each response is classified into 0 or more categories

# Analysis Stream in IBM SPSS Modeler



- Connects text and cluster analyses
- Generates output for feedback report
- Allows rapid analysis of new data sets

# Feedback Report



- Gene expression
- Cell function and physiology
- Different DNA

	<b>Cluster 1 44%</b>	<b>Cluster 2 28%</b>	<b>Cluster 3 27%</b>
<b>Cluster description</b>	Gene expression	Cell function and physiology	Different DNA
<b>Sample Response</b>	Each cell has different genes turned on and off which expresses different proteins.	The function of the cells within the heart cells and eye cells is what makes them different.	Cells have different jobs Heart cells and eye cells are different because their DNA is different the DNA is coded into different RNA strands which code for proteins that are very different...

# Distribution of Categories by Cluster

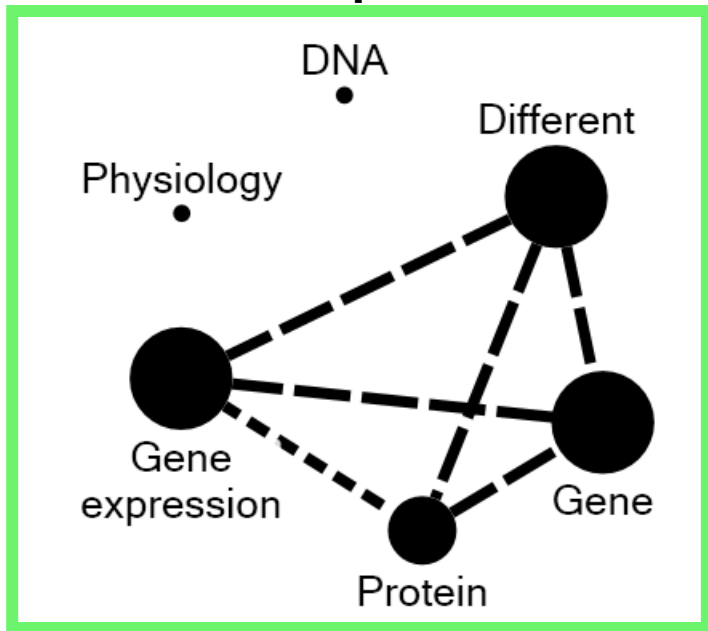
	Cluster 1	Cluster 2	Cluster 3
Lexical category	Gene expression	Cell function and physiology	Different DNA
Genes	98%	6%	11%
Gene expression	79%	2%	2%
DNA	15%	14%	88%
Physiology	11%	66%	17%
Protein	60%	10%	46%
Different	92%	75%	94%

>70% of responses were assigned to the category

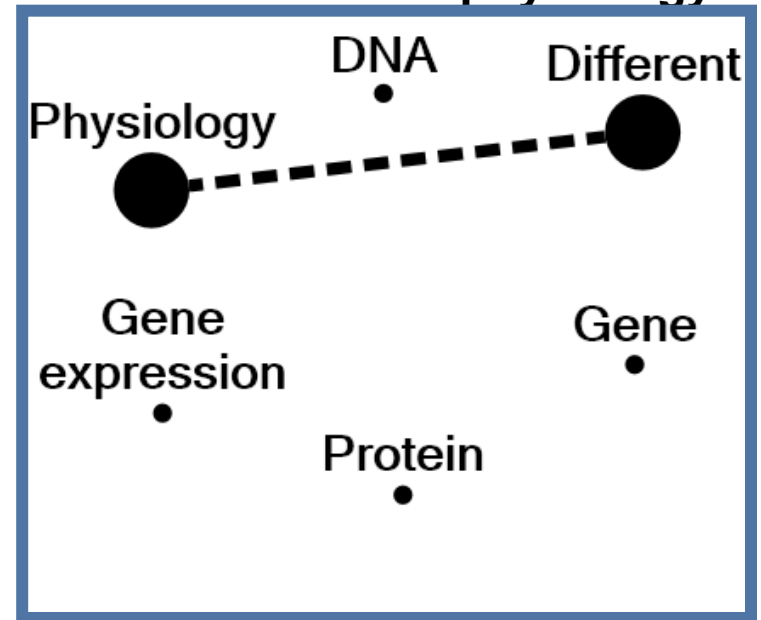
30-70% of responses were assigned to the category

<30% of responses were assigned to the category

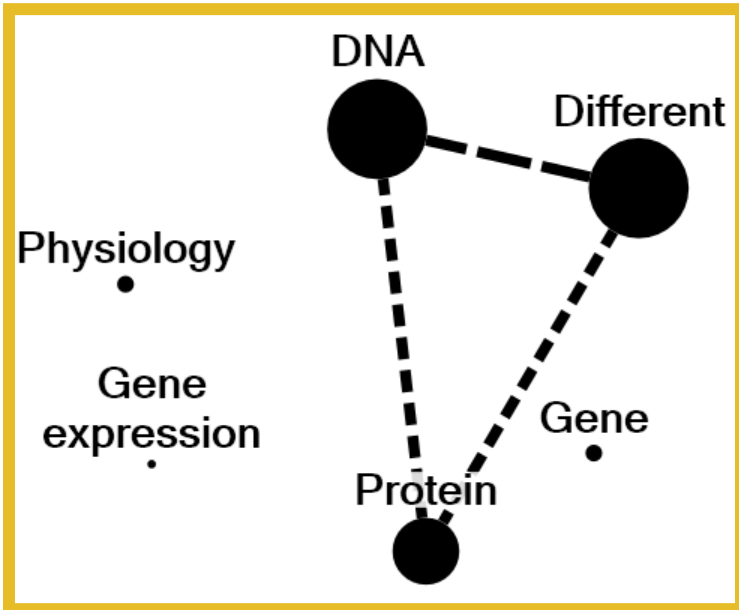
## Gene expression



## Cell function and physiology



## Different DNA



Node size: proportion of responses in category

- Share 75 – 100 % of responses
- - - - - Share 50 - 74% of responses
- - - - - Share 25 – 49% of responses

# Comparison of Student Responses PRE/POST Instruction

		Post-instruction			
		Final distribution	44%	28%	27%
Pre- instruction	Initial distribution	Cluster	Gene expression	Cell function and physiology	Different DNA
	35%	Gene expression	70%	13%	17%
	37%	Cell function and physiology	23%	51%	26%
	28%	Different DNA	41%	18%	42%



# Instructors response to CR questions and feedback reports

- Created clicker questions and led discussions based on results from feedback report
- Reported that written assessments were particularly important for gaining insight as to *why* students have struggled continuously with certain concepts
- Proposed future in-class activities to improve student writing skills

# Improving Automated Analysis for JiTT

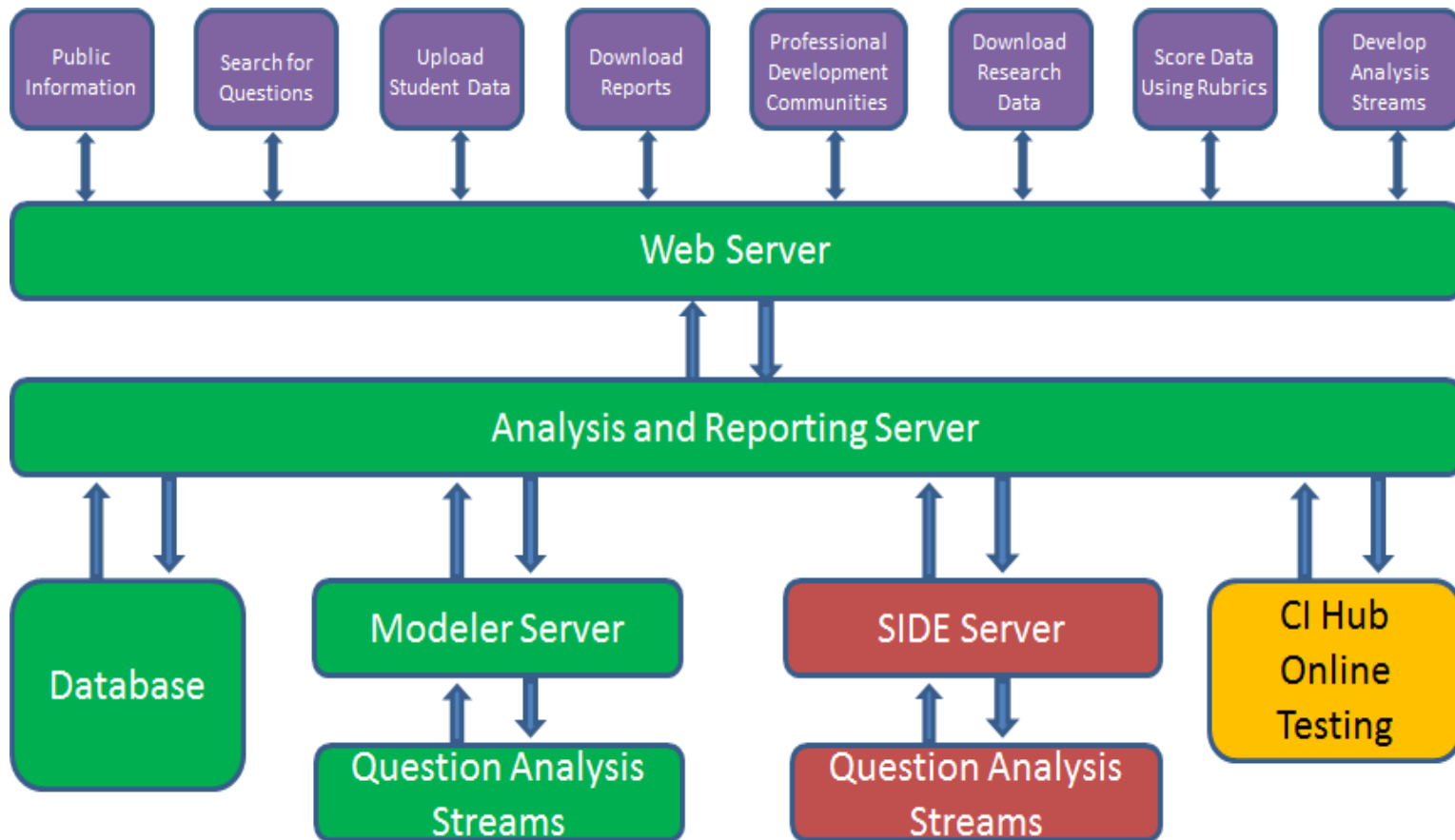
- Encourage student participation by giving credit for homework assignments
- Allow more time between each assignment and the next class for preparing instructional activities
- Professional development for faculty to help them address concepts that students find challenging

# Future Directions

- Faculty Learning Communities
  - Local: groups of faculty within a department or teaching the same course
  - Virtual: faculty across institutions
  - Use the same assessment & share instructional materials

# Future Directions

- Web Portal
  - support rapid assessment and feedback



# Acknowledgements

- Automated Analysis of Constructed Response Research Group (AACR)
  - Michigan State University
  - University of Colorado – Boulder
  - University of Maine
  - SUNY – Stony Brook
  - University of Georgia
  - University of California – LA
  - Western Michigan University



This work was supported by the National Science Foundation (Grant 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 National Science Foundation.