Moving Across Scales:
Using Lexical Analysis to Reveal Student Reasoning about Photosynthesis

Casey Lyons, Shauna Jones, John Merrill*, Mark Urban-Lurain, Kevin Haudek
*Corresponding author; Biological Sciences Program, 218 N Kedzie Laboratory
Michigan State University, East Lansing, MI 48824

1.1 Problem

In large enrollment STEM courses, it can be challenging for an instructor to gain adequate feedback from students in the classroom, thus inhibiting the instructor's ability to gauge student understanding of course material. Limited time and resources in large enrollment STEM courses means that multiple choice examinations are the primary assessment in these courses. Multiple choice assessments can be misleading (Nehm & Schonfeld, 2008), suggesting greater student understanding than is actually present. By contrast, student writing about biological processes reveals nuances about their thinking that cannot be obtained using closed-form assessments (Birenbaum & Tatsouka, 1987). It would be highly beneficial to both the instructor's teaching and the students' learning if more informative methods of feedback could be used in large enrollment courses. For instructors to move towards assessing students' principled reasoning, it must become more practical to analyze constructed (open-ended) responses in large enrollment undergraduate courses. At Michigan State University, we have successfully used computerized lexical analysis to study open ended assessments to gauge student's understanding in biology, chemistry and geology (Urban-Lurain et al, 2009). Computerized lexical analysis opens the possibility of evaluating large numbers of open-ended responses, which can provide instructors insight into patterns in undergraduate's reasoning.

1.2 Research Question

We were interested in two research questions:

1) How can we better reveal and understand students' complex ideas of photosynthesis?
2) Can we use lexical analysis software, which allows the processing of large numbers of student responses, to reveal common patterns of ideas about photosynthesis?

In this study, we examined the way that undergraduate students in an introductory cells and molecular biology course think about photosynthesis. We investigated photosynthesis because it is a complex biological process that requires understanding several key biological concepts: energy transformations, molecular rearrangements, and structure/function relationships. These are all "big ideas" that previous research has documented students struggle to understand (Wilson, 2006).

We used a crossover design, comparing how students answered multiple-choice and constructed-response questions on an exam. We used computerized lexical analysis of the constructed-response items to help us characterize student writing about photosynthesis. We wanted to determine whether students tend to have a clear correct understanding of photosynthesis -- a homogeneous set of ideas -- or if their answers involved many concepts, right and wrong -- a heterogeneous set of ideas.

2.0 Methods

We used computerized lexical analysis to help us examine the data we collect from student's open ended responses. SPSS Text Analysis for Surveys 3 (TAFS) (SPSS, 2009) is commercially available, lexical analysis software that we use to analyze the responses. Using computerized lexical analysis allows us the opportunity to build upon resources from previous lexical analysis projects, developed by our colleagues. Furthermore, we will be able to continue to assess the advantages and limitations of lexical analysis. This may help determine if automated analysis can be a useful tool in allowing instructors to utilize open ended questions in large enrollment undergraduate courses.

2.1 Data

We administered one of two possible essay and multiple choice questions to each student in an introductory biology course (Introduction to Cells and Molecules) at Michigan State University during a midterm exam during fall semester, 2009 (n=391). Each student received either the corn or maple tree question as multiple choice, and the other question prompt version as an open-response question. For example, if a student received the corn seed version of the multiple choice (#2 in 2.1.1), they received the maple tree version of the open-response (#2 in 2.1.2). By asking students to answer similar questions about photosynthesis in two different assessment formats, we are able to investigate differences in student outcomes. The crossover design allowed us to determine if the surface features of the questions (corn vs. maple tree) altered student responses.

2.1.1 Multiple Choice Questions

Students were given one of the following two multiple choice question stems. The answer choices were the same for both stems. Both questions were originally developed through research into how students trace matter through biological processes (Wilson, et al, 2006).

1. A mature maple tree can have a mass of 1 ton or more (dry biomass, after removing water), yet it starts from a seed that weighs less than 1 gram. Which of the following contributes most to this huge increase in biomass?
2. Each Spring, farmers plant about 5-10kg of dry seed corn per acre for commercial corn production. By the Fall, this same acre of corn will yield approximately 4-5 metric tons harvested corn. Which of the following processes contributes the most to his huge increase in biomass?

Both questions used the same set of answers (the correct answer, C, is indicated in italics).

A. Absorption of mineral substances from roots
B. Absorption of org. substances from soil via roots
C. Incorporation of CO2 gas from atmosphere into molecules by green leaves.

D. Incorporation of H_2O from soil into molecules by green leaves
E. Absorption of solar radiation into the leaf

2.1.2 Open Ended Response Questions

One of the following constructed response versions was given to each student in a crossover design intended to reduce the cueing effect. For example a student who received the maple tree version in multiple choice was given the corn version in constructed response.

1. Each Spring, farmers plant about 5-10 kg of dry seed corn per acre for commercial corn production. By the Fall, this same acre of corn will yield approximately 4-5 metric tons of dry harvested corn. Explain this huge increase in biomass.
2. A mature maple tree can have a mass of 1 ton or more (dry biomass, after removing the water), yet it starts from a seed that weighs less than 1 gram. Explain this huge increase in mass.

Both questions require the student to follow the movement of mass through the metabolic process of photosynthesis. All essay responses gathered from the exam were transcribed into a spreadsheet and then imported into SPSS Text Analytics for Surveys.

2.2 Lexical Analysis Libraries

TAFS software extracts *terms* (individual words or phrases), from open-ended responses based on *libraries* containing possible words or phrases. The software comes with a number of libraries and allows users to create customized term libraries. We have created a number of custom libraries that contain terms relevant to biology. The custom library for this project contains a variety of terms appropriate to cellular metabolism (Moscarella et al 2008). The library contains words that are appropriate to the topic, based on 1) an ideal response an instructor would expect from students to the questions posed and 2) terms that appear frequently in student responses.

2.3 Lexical Analysis Categories

The software provides a variety of automated semantic and manual methods of classifying the terms into larger conceptual groupings called *categories*. Our goal was to make the categories meaningful enough to provide information to instructors about the responses placed in them as well as to be encompassing enough to categorize nearly all of the responses. To do this, we created the categories based on an ideal response to the question and also attempting to characterize common student misconceptions. This way the instructor will not only have an idea of how well the students are grasping the material, but also an idea of common student misunderstandings. For instance, the category *Any Mention of Carbon Dioxide* was intended to capture explanations from students that provided a correct response. For a student's response to be included in this category, the response would have to contain the word "carbon dioxide", "CO2", or variations involving spelling errors. Categories were also designed to capture incorrect ideas as well. The category *Substances* attempts to capture explanations from students that discussed nutrients and minerals from the soil as the source of biomass. While this category would not be helpful in determining an accurate answer, it does help to highlight a common misconception students seem to have when following matter through autotrophs. We created a total of 16 categories (Fig. 1) to analyze this question. The categories are not mutually exclusive; each answer typically is placed into more than one category (mean=5.0, std dev = 2.6). The software categorized 99.5% of 391 responses. The remaining responses are considered "uncategorized" and contained only irrelevant information (e.g., "The huge increase in biomass is due to" and "The increase in biomass is due to tree gradually growing over the years. As it grows, the trunk gets larger, producing more mass and the branches extend and grow in length and diameter, also increasing mass.")

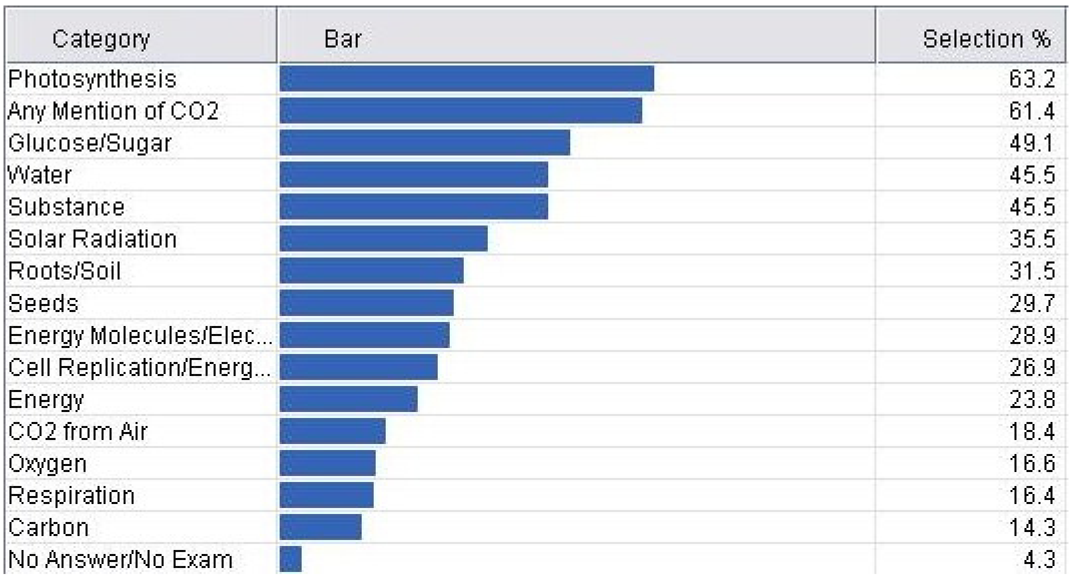

Figure 1. Frequency of categories in students constructed responses. Since categories are not mutually exclusive and multiple categories can be used in any single student response, the total exceeds 100%.

3 Results

3.1 Multiple Choice Data

In the crossover design, each student received one of two versions of the multiple choice question (corn or maple; section 2.1.1 above). A comparison of response distributions between these two versions demonstrates no significant differences (ANOVA, p=0.99), and so further analyses are based on pooled data. The answer in italics, C, is the correct answer and was chosen by the majority of students, nearly 60%.

A. Absorption of mineral substances from roots (7.7%)

B. Absorption of organic substances from soil via roots (12.7%)
C. Incorporation of CO_2 gas from atmosphere into molecules by green leaves (59.4%)
D. Incorporation of H_2O from soil into molecules by green leaves (7.7%)
E. Absorption of solar radiation into the leaf (12.7%)

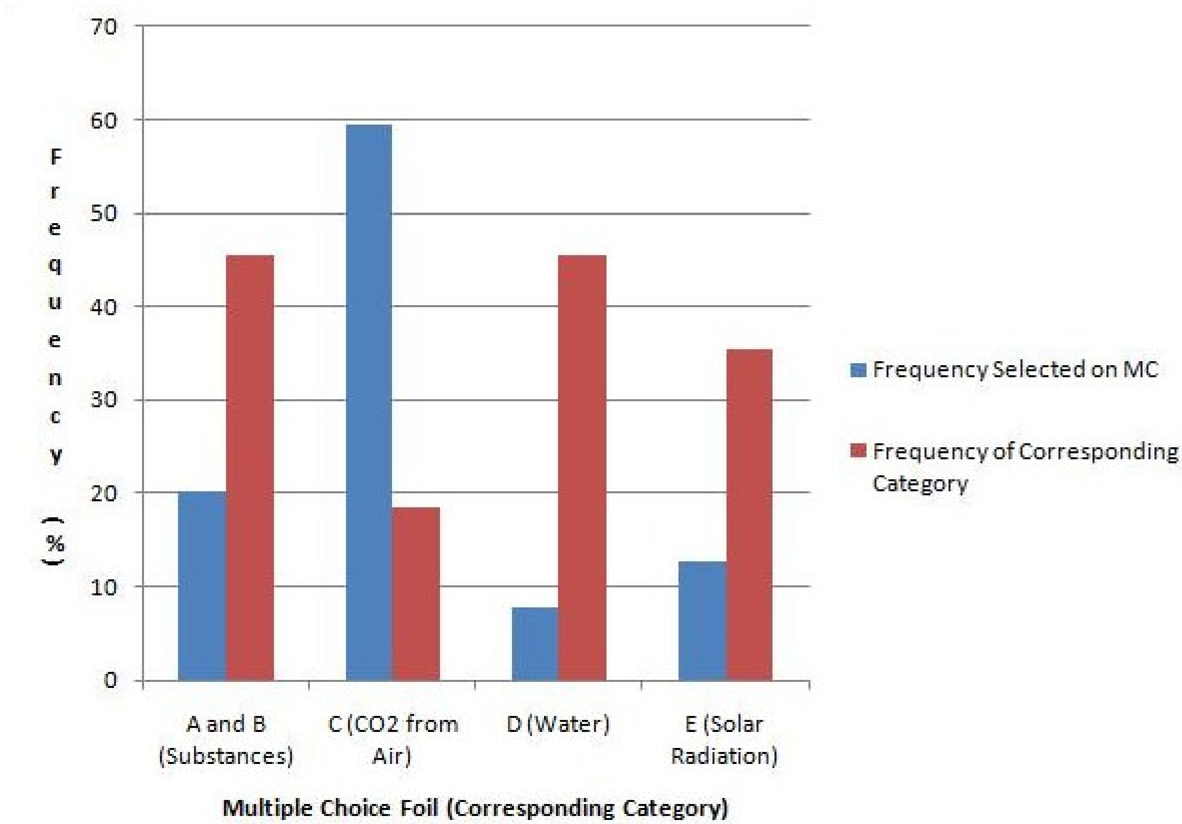

Figure 2. Comparison of choice selection frequency and written concept frequency. Blue bars are frequencies of multiple choice answer selection. Multiple choice answers A and B are combined in this figure because of conceptual overlap "substances from roots" in the distractors. Red bars are frequencies of the corresponding lexical analysis categories (indicated in parenthesis) observed in all constructed responses. Because students can include more than one category in a written response, the total exceeds 100%.

Common incorrect concepts about photosynthesis are expressed by the distractors in the multiple choice version of the question. The ideas present in the multiple choice distractors are captured by categories that are often associated with misconceptions about photosynthesis. A comparison of the lexical analysis data and multiple choice data (Fig. 2) clearly show that while 60% can <u>choose</u> the correct answer, less than 20% clearly identified the related concept in their writing. (It should be noted that the multiple choice and constructed response questions occurred on the same page of the exam, separated by 6 items, so students had already been cued on this topic during the exam.) When constrained to select only one answer, only 20% of all students select either A or B (substances taken up by roots), whereas approximately 45% of students include the idea in their constructed responses (the first pair of bars). The frequencies are higher for each corresponding category than the frequencies for the related multiple choice answers. The differences between the frequencies of students selecting a multiple choice answer and the frequency of students falling into that corresponding category also highlight the heterogeneity in student thought that is not captured by the multiple choice.

3.2 Student's Heterogeneous Ideas about Photosynthesis

Comparing the results from the open-ended and multiple choice questions provides different insights into students' understanding of photosynthesis. Almost 60% of students answered the multiple choice question correctly, but that number may be deceiving compared to the constructed responses. Furthermore, many students who received credit for the multiple choice question in fact had heterogeneous ideas: a combination of correct and incorrect concepts in their open-ended responses.

The first example response, below, comes from a student whose multiple choice selection was correct.

> *The increase of biomass is due to the concept of photosynthesis. Most of this process is fueled by the nutrients that are absorbed in the soil. They are used to produce ATP and energy via the plant root. As the plant grows, photosynthesis continues to happen through the light reactions and Calvin cycle, O2 and light are absorbed and ATP and NADPH are produced and used for the Calvin cycle. This process uses glucose to convert it to CO2. ADP and NADP are then recycled back to the light reactions to produce more ATP.*

Here we see that the student appears to be struggling to put together the various complexities of photosynthesis describe above (section 1.2). Note, for example, the incorrect attempt to describe soil nutrients providing "fuel" for the process, production of ATP and energy in the plant root, involvement of O2 being used by photosynthesis, and confusion with cellular respiration (converting glucose to CO2). Many of these words are indeed used in an expert explanation, but this student lacks an accurate narrative ability to put them together in a coherent explanation. This is in marked contrast to the ability to select the correct multiple choice answer.

Heterogeneity also applies to students who selected the incorrect multiple choice answer. The following example shows the response of a student whose multiple choice selection was incorrect (B. Organic material from the soil), yet the student's writing expresses some correct ideas.

> *The huge increase in biomass comes from the plants carrying out photosynthesis. This causes them to produce energy for themselves, which in turn, makes them grow. The plants absorb organic materials from the ground and CO2 from the air which both contribute.*

Note that the student indicates some understanding that the plant's biomass comes from carbon dioxide and photosynthesis. Additionally, however, the student mixes in some incorrect concepts, for example that the mass originates as organic materials from the ground. The combination of a biologically accurate explanation of an increase in biomass and the common misconception that plants gain mass through organic materials from the ground demonstrates that correct and incorrect ideas can coexist. Recall that this student made an incorrect multiple choice selection, and yet in writing has shown ideas that are at least partially correct. Our goal is to accurately diagnose these cases of mixed conceptions in order to better target instruction.

3.2.1 Representing heterogeneity of ideas

We are interested in how students may express the correct idea in combination with other concepts. One powerful way to visualize this is via a semantic network (Steyvers and Tenenbaum, 2005), which shows the co-occurrence of categories within students' responses (Figure 3). One category representative of a correct idea is captured by the category *CO2 from Air* (72 of 391 responses) and therefore we have selected this as the base category for the figure. In order to avoid the clutter created by low frequency co-occurrences, in this diagram we have set a threshold minimum of 25% co-occurrence (18 shared responses). In other words, at least 18 students must have used another category in combination with *CO2 from Air* to be included as a node. For example, the category *Substances,* was observed in 32 of the *CO2 from Air* responses (44%). Other examples include *Roots and Soil* (33%), and *Energy Molecules/Electron Trans Molecules* (29%). Or, stated differently, 44%, 33%, and 29% of students, respectively, when expressing the "correct" idea that the mass comes from CO2 uptake from air, <u>also</u> mention other sources as contributing to the biomass, suggesting that they are expressing multiple ideas.

From this analysis we can see that significant numbers of students invoke multiple concepts in their explanations. Some of these concepts may be thought of as "correct" in our scientific context (e.g. *CO2 from Air*) while at the same time their explanations also contained concepts that would be classified as "incorrect" (*Roots/Soil*) in the same context. When we analyze the constructed responses, then, we reveal a rich mix of ideas expressed by our students.

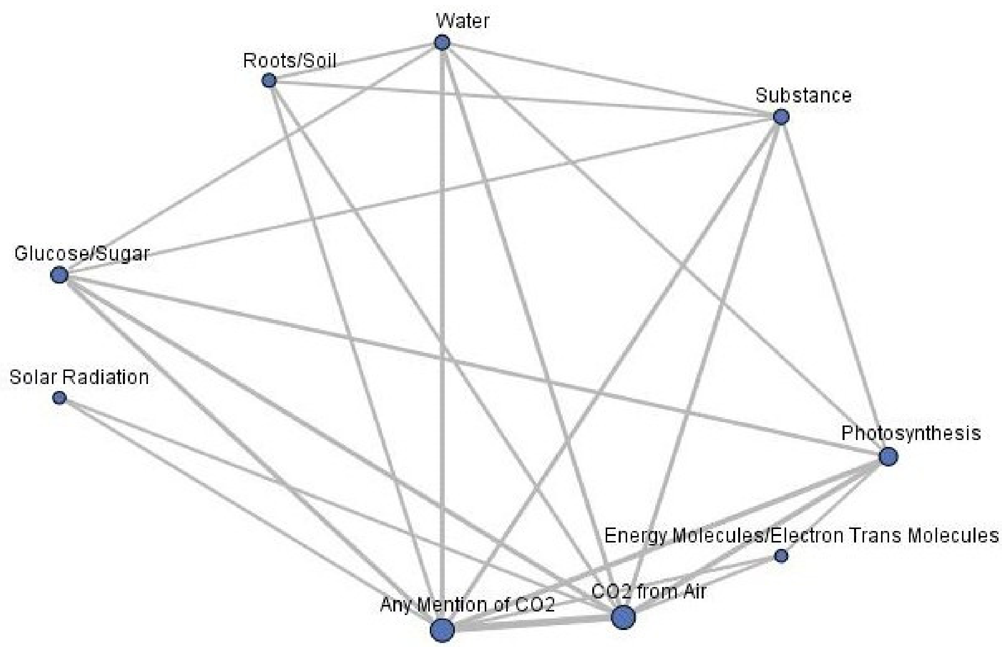

Figure 3. A semantic network diagram depicting conceptual connections in student responses that included “CO2 from air”. Categories used by at least 18 students (25%) are included, size of node and thickness of line indicating relative frequency of co-occurrence.

3.3 Relationship between Categories and Multiple Choice

Another way to examine the differences between the multiple choice and constructed response information is to compare students’ multiple choice selection to the concepts employed in their written response (Fig. 4). The two categories most commonly associated with correctness and incorrectness, *CO2 from Air* and *Substances* respectively, show major differences in frequency between students who answered the multiple choice version correctly. As would be expected, students selecting the correct answer for the multiple choice were much more likely to mention carbon dioxide absorption from the air as those who selected one of the incorrect distractors. Conversely, vague references to substances in the soil, as captured by the categories *Substances* and *Roots/Soil*, were much more likely to be in the constructed response answers for students who answered the multiple choice incorrectly.

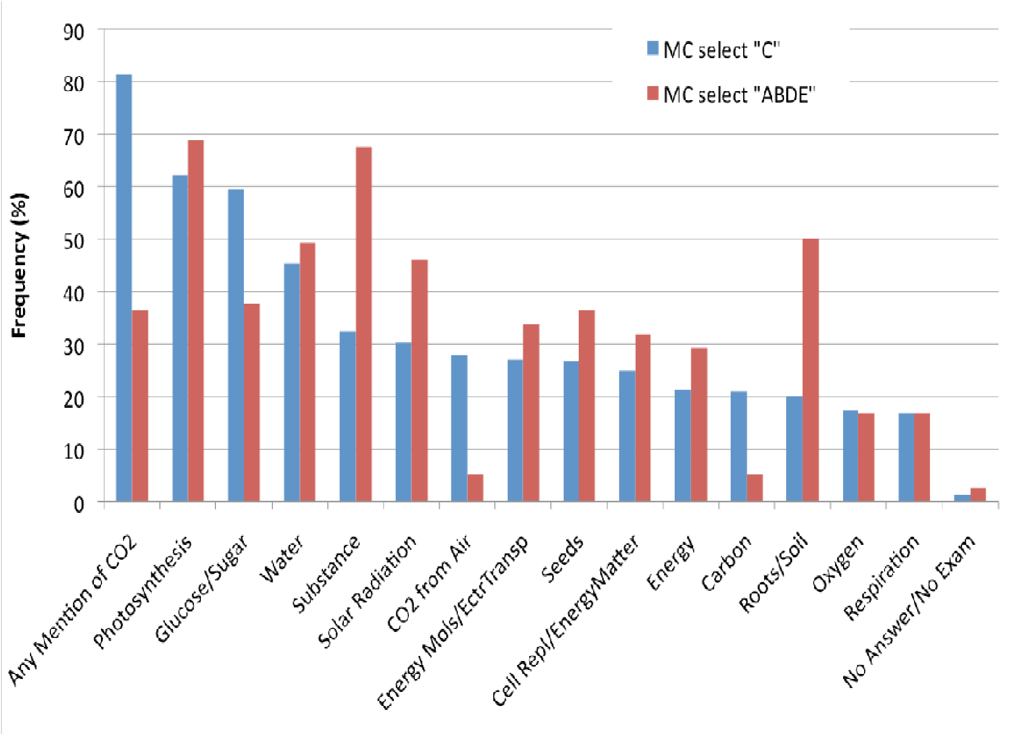

Figure 4. Categorization of the student responses broken down by the students' multiple choice selection. Categories generated in lexical analysis are listed along the bottom. Constructed responses written by students that selected the correct multiple choice answer (C) are indicated in blue. Responses written by students that selected an incorrect multiple choice answer (A, B, D or E) are indicated in red.

4.0 Discussion

Carefully designed multiple choice questions are normally accepted as a best case option for formative assessment in large enrollment classrooms where the use of constructed response assessment items are impractical. The value of these items is often enhanced when the distractors are derived from common student errors or misconceptions. Even though students may harbor these misconceptions, however, their prevalence may be masked by the students' ability to select the one "correct" response from the choices. We have shown how computerized lexical analysis can be used to reveal the more complex mix of student ideas even in extremely large enrollment courses. With this information, instructors have a far more complete understanding of their students' ideas, and can more precisely target classroom interventions.

The primary strength of this technique we believe to be for formative assessment for guiding instruction. Although one might extract quite useful information from simply reading a subset of students' responses, it is highly unlikely that such a sampling would accurately reveal the patterns in responses that can be generated via the techniques that we have employed. We believe that the technique may also offer promise for student-formative assessment, although we have not yet attempted this extension. We envision a system in which online homework problems are analyzed in real time with feedback given as study guidance to the student based on the combinations of ideas found in the submission.

Computerized lexical analysis is also an excellent research tool. With automated analysis, constructed response assessments can be used to evaluate pedagogical innovations with a much reduced demand on resources, and with the accuracy provided by analyzing every response rather than a sampling.

We are currently extending this work in several areas of biology, including genetics, natural selection, and cellular metabolism, with the goal of creating complete constructed response assessment instruments that could be applied in combination with or in place of concept inventories. We are hopeful that this research will yield valuable tools that will make the use of constructed response far more accessible especially to instructors in large enrollment courses where this form of assessment is otherwise seldom used.

REFERENCES

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

Moscarella RA, Urban-Lurain M, Merritt B, Long T, Richmond G, Merrill J, et al. Understanding undergraduate students' conceptions in science: Using lexical analysis software to analyze students' constructed responses in biology. Baltimore, MD: National Association for Research in Science Teaching; 2008.

SPSS. (2009). *SPSS Text analysis for surveys 3 user's guide*. Chicago, IL: SPSS, Inc.

Steyvers, M., & Tenenbaum, J. B. (2005). The large-scale structure of semantic networks: Statistical analyses and a model of semantic growth. *Cognitive Science*, 29(1), 41-78.

Nehm, R.H. and Schonfeld, I. (2008). Measuring knowledge of natural selection: a comparison of the CINS, and open-response instrument, and oral interview. *Journal of Research in Science Teaching*, 45(10): 1131-1160.

Urban-Lurain M, Moscarella RA, Haudek KC, Giese E, Sibley DF, Merrill JE. Beyond multiple choice exams: Using computerized lexical analysis to understand students' conceptual reasoning in STEM disciplines. Vol 39. San Antonio, TX: ASEE/IEEE ; 2009.

Wilson C.D., Anderson C.W., Heidemann M., Merrill J.E., Merritt B.W., Richmond G., Sibley D.F., Parker J.M. (2006) Assessing students' ability to trace matter in dynamic systems in cell biology. CBE Life Sci Educ 5:323-31.