

## Building the Biology Concept Inventory

University of Colorado, Boulder

Kathy Garvin-Doxas

Center for Integrated Plasma Studies

Michael Klymkowsky

Dept. of Cellular, Molecular, & Developmental Biology

About 5 years ago, Mike developed a course called BioFundamentals. The course was designed around virtual laboratories, over the years, he integrated interactive engagement approaches to teaching and learning into his lectures and interactions with students. As the course evolved, he realized that he didn't have a thermometer – in other words, he was building a new and innovative course at the introductory level of biology, but he had no reliable way to ensure that his students were learning at the level of conceptual understanding that he intended. Were his students integrating the knowledge they were exposed to at the analysis, synthesis and evaluative levels of the taxonomy Bloom and his colleagues proposed in 1956? Were they even learning at the application level? With the help of a small grant from Mike Zeilik's group (FLAG), Mike and I began exploring what was available at the time in terms of learning assessments in the biological sciences and the research that had been conducted on student conceptual understanding (this was in 2003). Eventually, we proposed the project we are currently working on, Building the Biology Concept Inventory (BCI), funded by NSF.

**Goals and Objectives:** The BCI project goal is to design a 20-30 item CI instrument that is capable of measuring student understanding of fundamental biological concepts at the undergraduate level, can serve a large population of introductory courses; and can be employed to monitor the effectiveness of teaching innovations with respect to those fundamental concepts. In order to develop this instrument, we needed to:

1. Identify students' commonly held misunderstandings or misconceptions at the introductory level
2. Create an instrument capable of accurately representing students' location on the conceptual landscape using language that forced them to step outside their habitual and internalized use of the rhetoric of science
3. Create a development process that is replicable and allows us as well as others to expand the instrument in order to offer greater coverage of the broad scope of concepts covered at the introductory level

Our hope is that by providing an instrument that can be used by a wide variety of programs in the biological sciences, the BCI will be seen and used as a reliable means of determining the relative success of teaching innovations introduced when compared to traditional lecture methods and thereby provide a meaningful measure that could serve as a catalyst for continued improvement in teaching in the biological sciences, much like the Force Concept Inventory (FCI) has for physics.

Our instrument is designed meet the needs of several groups involved in undergraduate education. This CI is designed to be useful/helpful for those who are teaching at the undergraduate level. One of our long-term goals is to create something that students can use to track their own progress in understanding of the biological sciences and that can eventually help them identify material they need to revisit (either through on-line tutorials or additional reading). The implications of the instrument should engage the research community in biology education to further explore any areas of commonly held misunderstandings and/or misconceptions that have not previously been identified, so that appropriate teaching interventions can be created and we can better understand how best to help a wide variety of learners master their understanding of fundamental areas in the biological sciences.

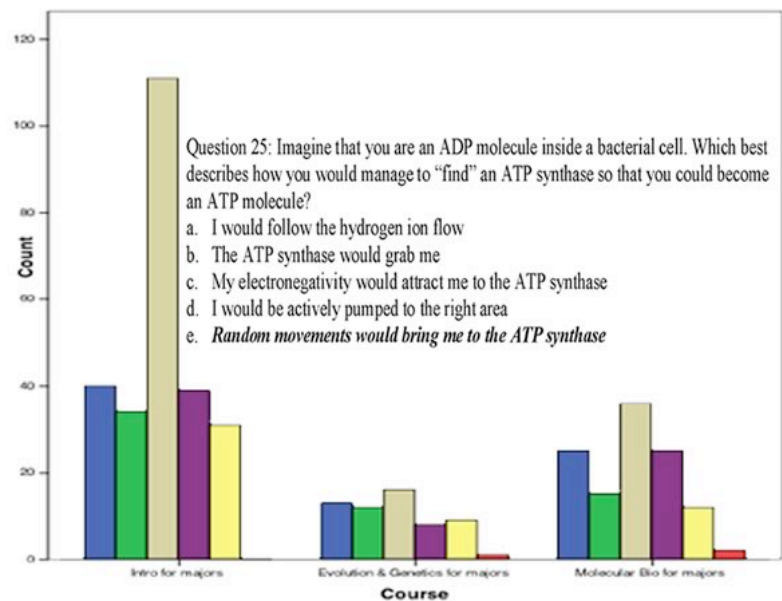
The BCI is in a multiple-choice format, but distracters are based on commonly held student misconceptions identified through research. The instrument is useful when administered prior to instruction as a means of allowing professors to better understand the needs of students in their particular course (what areas the majority of students have already mastered, areas where they are not as advanced in terms of meaningful understanding, etc.). It can also be used for comparison when pre-

instruction results are compared with post-instruction, so that professors can determine the degree to which their teaching approach in a particular conceptual area has improved student understanding and those areas where they need to reconsider their approach. It is designed to measure understanding in naturally occurring clusters or groupings of related material. Results of administering the BCI to students at the introductory level and to those at the senior level, indicate that the instrument has potential for use as part of over-all program evaluation (statistical results and interviews with students indicate the instrument is capable of demonstrating whether or not advanced students have moved further along in terms of overall understanding of fundamental principles in biology when compared to students at the introductory level).

Currently, the BCI is limited and provides partial coverage of only three overlapping clusters of conceptual understanding:

1. Biological Processes: the interrelationship between evolutionary processes, genetic traits, and the molecular nature of mutation
2. Biological Structures: the relationships and interactions among structures such as molecules, DNA, lipids, and membranes
3. The Idea of Random: this cluster overlaps with genetic processes, in particular genetic drift, and includes specific questions about diffusion.

We have administered our alpha version (n=539) in 2005, our beta version (n=442) in 2006, and our pilot version (3 participating universities; n~500) in 2007. Both the alpha and beta versions were administered pre-instruction, but modified for post-instruction administration, and the pilot version has been administered pre-instruction and will be administered again at the end of this semester. We are in the process of making final revisions to the instrument in terms of wording, questions that are included, etc. based on results of think-aloud interviews with students (n=20 for Fall, 2006; n=40 for current number of project interviews). For the development process, structured think-aloud interviews (described below) covered the entire instrument, but we used results from comparisons of introductory and advanced students on individual items for each the alpha and beta version as a means of identifying questions and responses that required further exploration (either validation, elimination, or rewording).



In the figure, we show a comparison of Fall 2006 responses taken from an introductory course, a third semester course, and an upper division course. Results are reported in numbers of students who selected each response, rather than percent of the number of students in the course who selected each response (red=number of missing responses). The question and responses were a result of small group classroom discussion and third-round essay question coding. Each potential response represents a recurring pattern in student explanations of the process. Based on item analysis alone, it would be easy to discount and discard this question or

elect to change these potential responses. However, several other sources of qualitative data indicate that this is actually a very good question and it does represent a fundamental and persistent misconception that students have about diffusion and other random processes. During small group discussions with her students enrolled in an introductory course, one of our student assistants from the Noyce Scholars

Program (Rachel Gheen), realized that they experienced difficulties with this process. They could list and label what happens (at the rote level), but they failed to understand how and why the movement occurs. We developed several essay questions to further explore these concepts and our content analysis led directly to this question with these distracters and correct response. When we followed through with structured think-aloud interviews, we found that students interpreted the question in the way we intended; that they attributed the meaning we intended to each potential response; and that the response they selected represented their true understanding of how and why the movement happens during this process. More importantly, their struggle with this concept is consistent with the difficulties traced through results of our other questions relate to the “Idea of Random.” We continue to follow student misconceptions about diffusion and other random processes in interviews and additional essays and analyses. Thus far, our think-aloud interviews indicate that one reason students can memorize rote information about diffusion and other random processes, but experience extreme difficulty with these ideas at the conceptual level is because so much of their learning experience in the biological sciences stresses the “efficiency” of biological systems and “random” processes are not seen and experienced as “efficient” by students. Students may experience them as mutually exclusive from one another, perhaps because they confound the vernacular use and meaning of “random” with what random means in science (note: the only student in at the introductory level who was capable of discussing “random” and diffusion in the sense it is used in the sciences, was a theoretical mathematics major who claimed that he learned about these processes in his physics courses and the study of quantum mechanics).

Obviously, we’ve jumped ahead here during this discussion of the figure, its interpretation, and how we pursued the results through triangulated methods, but we would also like to share more about the actual process we have developed and follow in construction of the BCI.

**Development and Process:** Our development process employs a combination of qualitative and quantitative data collection and analysis techniques. It is an iterative process that is heavily focused on understanding misconceptions through student discourse. We collect this discourse in the form of essays and thematic interviews as a means of, identifying what students elect to discuss (no matter what question is posed); the connections they make among ideas; and how they talk about fundamental ideas in biology.

*Essays and Coding:* We administered a series of general, over-arching essay questions to thousands of students and asked them to write at least 100 words in response. From the perspective of content analysis, essays were coded based on the question: what recurring patterns can be found not only in students’ responses to individual questions, but to the questions in general? Content analysis has been used in the social sciences since the 1940’s and can be described as “any technique for making inferences by objectively and systematically identifying specified characteristics of messages” (Holsti, 1969). By the 1950’s, content analytic approaches moved beyond a focus on simply counting key words, to included examination of relationships articulated or implied by the way words and phrases are consistently used, as well as oppositions between words, phrases, and ideas (e.g., deSola Pool, 1959). Essentially, content analytic techniques involve multiple coders reviewing data numerous times. Our analytic units include not only key words, but phrases, whole paragraphs, etc. taken in context, that allow us to identify recurring patterns in the data, which in turn, enables that data to be understood in terms of categories or groupings of words and phrases with similar meaning or *connotations* (Weber, 1990). Our work is qualitative in the sense that we employ emergent categories rather than *a priori* coding groups. Essentially, we let the students tell us what the patterns are rather than assuming categories of responses before coding. (An example of an assumed category would be looking for evidence that students confound “dominant” as “stronger than”.)

Based on the recurring patterns both within and across student essays, we developed a series of candidate questions along with candidate responses taken from student language. We began with a series of questions and potential responses that examined student understanding of *evolution* and *natural selection* as well as *genes and traits* and the notion of *dominant and recessive* (often cutting and pasting student language directly into the sample instrument). Potential responses were taken from the most frequently recurring patterns of response found in the content analysis. This is the process we used to identify areas and expressions of students' commonly held misconceptions.

*Thematic Think-aloud Interviews:* We conducted initial thematic think-aloud interviews with groups of students who had completed the introductory level course. Thematic think-aloud interviews start with the student talking through his/her response to and thoughts about the questions and potential responses as they read them. Discussion can be detailed and the thinking process is captured in addition to student suggestions for wording or even alternate questions. The student interviewee takes an active role in question development during this phase. Initially, we conducted 7 of these detailed interviews and came away with a better understanding of how student's discourse reflects their understanding. We were greatly assisted in this by the students themselves who demonstrated a remarkable ability to verbalize their thinking processes.

Again, since we began by asking, "what is?", we engaged in a qualitative and iterative process of furthering our question development by using better essay questions followed by additional content analysis and then thematic think-aloud interviews. The more we came to understand how students talked about biology, the better able we were to develop essay questions that were meaningful to them. These in turn, led to more facile (and directly connected) concept question development. Using this process, we developed an alpha version of the BCI and administered it to 539 students (at the introductory, middle, and advanced levels). The process described above continued, but we added another step to begin validation of alpha version questions.

*Structured Think-aloud Interviews:* In these interviews, students take the current version of the BCI instrument and, once they have completed it, they talk the interviewer through their interpretation of each question and how they selected their response. The process is much more structured than the more exploratory, thematic think-aloud interviews. This type of interview ensures that students are interpreting the questions in the way they were intended and that the answer they select represents their true conceptual understanding at that time. After results of these interviews are analyzed, improvements to wording are made, questions may be discarded, new essay questions are developed, and the iterative cycle continues until a complete instrument is developed. At that point, structured think-aloud interviews are focused on validity and reliability studies can be made.

Within this context of development, we implemented additional procedures in an effort to better communicate with one another and with potential users. The first procedure developed was a guide for question development:

1. Write a justification for each question that is grounded in the coding, interview results, and/or other rigorous research documented on the conceptual area in the biological sciences.
2. For each distracter and the correct response, indicate what selecting this response indicates about student understanding. Be sure to ground your statements in research results rather than personal opinion and/or experience. What is it that our research (or that of others) indicates that selecting this response reflects about student understanding as it relates to the goal and purpose of the question?

These are both areas that we refine as we collect and analyze more data through student essays and interview results. Here we offer an example of how our question development model works:

**25. Imagine that you are an ADP molecule inside a bacterial cell. Which best describes how you would manage to "find" an ATP synthase so that you could become an ATP molecule?**

*This question is designed to test whether students understand that diffusion is caused by random motion of molecules. Our research shows that most students do not understand that molecules are in constant motion. Although they understand the concept of gradients and moving down a concentration gradient, they don't understand what causes this movement.*

**a. I would follow the hydrogen ion flow.**

*Students who choose this answer think that ADP somehow can identify where a hydrogen ion gradient is. In our research, students never explain how an ADP would sense an H<sup>+</sup> gradient. Many students use language that suggests they think ADP is able to actively seek out an ATP synthase. For example, ADP is described as "looking for" or "noticing." Although the answer is incorrect, students who select may understand correctly that there is a hydrogen ion gradient. They may also correctly understand that the gradient is causing hydrogen ions to flow through the synthase and across the plasma membrane to the area where hydrogen ions are less concentrated.*

**b. The ATP synthase would grab me.**

*Students who select this answer think that an ATP synthase senses the presence of ADP and actively grabs it. Excluding some students who said ATP synthase has a receptor that recognizes ADP, most students do not explain how the ATP synthase does this grabbing. One again, students who select this answer may believe that molecules have the ability to actively seek out or choose other molecules.*

**c. My electronegativity would attract me to the ATP synthase.**

*Students who select this answer think that charges cause the ADP and ATP synthase to be attracted to each other. Based on what students know about molecular interactions, it's a good guess. However, it seems to be just that – a guess – because students don't say what charge the ATP synthase is or that the ATP synthase is oppositely charged.*

**d. I would be actively pumped to the right area.**

*Students who select this answer think that the ADP is somehow placed in the correct area so that it is close to the ATP synthase. Students may have no explanation for how this occurs.*

**e. Random movements would bring me to the ATP synthase.**

*This is the correct answer. In other words, ADP finds ATP synthase by the random motion of ADP molecules, which causes diffusion.*

*Refine, Reject, Add:* We compared the results of the alpha and beta BCI administration to structured think-aloud interview results and determined whether or not a question needs to be refined, rejected, or if perhaps a new question needs to be considered, explored and developed.

*Analyze Question Clusters:* On the one hand, we developed questions that explore different aspects of the same or similar phenomena in biology. On the other hand, the question clusters we developed did not occur by design and were grounded in student discourse. Once we developed a beta version, we examined it for patterns (content analysis again) and groupings that naturally occur – and we looked at the level of potential student responses. One of the areas we are currently working on is clearly defining the relationships and weights among questions and student responses. For example, in the above question, if a student selects response A, they do not fully understand the process, although they may understand that there is a hydrogen ion gradient, and that there is a net flow of hydrogen in the direction of the gradient. This question relates to other questions about processes, and the understanding of gradients here may be worthy of a "+" score in another cluster of understanding. These are the relationships within clusters and among questions that we continue to explore.

Currently, we are also working on reliability measures of our pilot instrument. We plan to ensure validity and reliability in terms of each cluster rather than the instrument as a whole, so that it can become more adaptable to the needs of those teaching. In addition, we continue to work on ensuring that the question model is clear, appropriate and rigorous for each question we retain.