Avoiding Reflex Responses: Strategies for Revealing Students' Conceptual Understanding in Biology

Michael W. Klymkowsky, Rachel Gheen, and Kathy Garvin-Doxas

The Bioliteracy Project, Molecular, Cellular, and Developmental Biology, University of Colorado, Boulder, Boulder, Colorado 80309-0347, U.S.A.

Abstract. There is widespread concern about the level of scientific literacy in the U. S. An important, although often overlooked, point, is that student learning is generally only a good as the assessments used to measure it. Unfortunately, most assessments measure recall and recognition rather than conceptual understanding, and as a result over-estimate levels of scientific literacy. We have encountered this fact during the construction of the Biology Concept Inventory (BCI). Using the concept of diffusion, which is taught in a wide range of introductory biology, chemistry, and physics courses, as an exemplar, we describe lessons learned and strategies we use to create questions that better probe student understanding.

Keywords: reflexive responses, memory versus conceptual assessments.
PACS: 01.40.-d, 01.40.gf, 01.50.Kw, 01.75.+m

INTRODUCTION

Educators have long realized that authentic assessments drive effective instruction [1]. Without such assessment instruments, instructors do not know whether their teaching has been effective, whether students have learned critical information, or whether student misconceptions remain intact, or have been created. At the same time, students and instructors can be content with ersatz understanding, which is often no more than the ability to recognize terms [2]. Assessments have begun to carry more weight as our country continues to emphasize the importance of developing a scientifically literate population and work force [3], as witness their role in the federal "No Child Left Behind" program.

Good assessments of science education determine students' level of subject mastery; they measure how well students understand fundamental scientific concepts and can use these concepts in problem solving or explanatory situations, that is, whether they can think and express themselves in a scientifically valid manner. Producing authentic tests is often substantially more difficult than it seems. However, given that assessments often determine how students are taught and what they learn, developing assessments that accurately reveal students' conceptual understanding is critical to attaining the goal of a scientifically literate citizenry.

Science educators have begun to develop a range of assessment instruments that focus on conceptual understanding rather than the recall of isolated bits of information. The Bioliteracy Project[1] models the development of a Biology Concept Inventory (BCI) after the Force Concept Inventory (FCI)[4, 5], which was developed for use in introductory college physics classes. Like the FCI, the BCI is a research-based, multiple-choice instrument designed for use in evaluating instructional methods and to assess students' understanding of key concepts. Through the development of the BCI, we are now finding for biology what the FCI found in mechanics: students taught through traditional methods are learning much less than we had assumed. Just as FCI findings served as a wake up call for many physics instructors [e.g., 6, 7], it is our hope that the BCI and similar instruments, such as the Natural Selection Concept Inventory [8] will lead to changes in biology course design and teaching methods.

BUILDING CONCEPT INVENTORIES

The first step in building any concept inventory is to discover what students believe about specific subjects. We began our project by asking

[1] See http://bioliteracy.net

students open-ended essay questions on topics often covered in high school and introductory biology classes. To analyze their responses, we developed a web-based Java program (Ed's Tools) that enables us to generate a database of student language that reveals at least part of their thinking in response to a specific question. Very quickly we recognized that certain questions, which appeared likely to be informative, were not. We will look at examples of poor questions and describe the strategies we use to transform poor questions into ones that better reveal students' conceptual assumptions. In describing our question-writing processes we will focus on the concept of diffusion. While this is certainly not the only conceptually difficult subject we have uncovered, it illustrates our question-writing process.

Poor Questions Elicit Reflex Responses

Diffusion is taught in nearly all high school and undergraduate entry-level biology courses, as well as in chemistry and physics courses. Understanding diffusion is critical for understanding how molecules move into and out of cells, as well as the basic mechanisms of molecular interactions. Diffusion can be understood at a number of levels. As recognized by Sutherland and Einstein, it is random molecular motions that drive diffusion. In the context of the biology classroom, diffusion is commonly presented in the context of concentration gradients and the tendency of a given molecule to move from high to low concentrations across a biological membrane. While net flux depends upon diffusion, diffusion often occurs without net flux. We can define those conditions in which a net flux of molecules will occur if we understand the basic random nature of molecular motion.

We began to probe students' understanding of the motion of molecules by asking a simple question: "What is diffusion and why does it occur?" After gathering ninety-seven student essay answers, we found that none of the students understood (or better put - none of the students explicitly stated) that diffusion is based on the random motion of molecules. Instead, we discovered that the question elicited what we call a "reflex response." Reflex responses are automatic answers that rely on recalling the context in which the questions subject was presented, either during the course of lecture or in assigned readings, rather than a response based on understanding the processes involved in actually answering the question. In the case of our initial diffusion question, essentially all of the students described a set of facts related to concentration gradients and membranes, rather than diffusion *per se*. Many students listed the types of molecules that could and couldn't diffuse. While many of these responses were accurate in and of themselves, they failed to answer (and generally ignored) the question posed. All of this points to the key feature of a reflex response: students memorize and supply information as a discrete package with little regard to what the question actually requests.

We have found that many standard science assessments contain a number of questions that elicit reflex responses; in some cases, they actually offer a prompt for what the correct reflex response should be. For example, consider a question from an eighth grade national science assessment (carried out by the Institute of Educational Statistics, US Department of Education and obtained from their website): "*Some people have proposed that ethyl alcohol (ethanol), which can be produced from corn, should be used in automobiles as a substitute for gasoline. Discuss two environmental impacts that could result from substituting ethyl alcohol for gasoline*." An example of a correct and complete student response is: "*We would need a lot more land, soil, and money to grow enough corn to feed people and to put in cars. We would have to cut down forests, causing to higher CO2 levels and making more animals endangered. We would need more irrigationing (sic), using up our small % drinkable water.*"

The students' answer makes a number of unwarranted assumptions, triggered apparently by the suggestion of using corn, and is a reflex in that it does not explicitly discuss ethanol versus gasoline, which the question asks, but rather corn alone. It fails to note that growing corn leads to the sequestration of CO_2, and so its burning produces a lower net increase in CO_2 levels compared to the combustion of fossil fuels, depending, of course, on the extent to which the generation of ethanol requires the use of fossil-fuels. It also assumes that forested lands are more productive in sequestering CO_2 than agricultural crops and it does not consider other sources for the generation of ethanol (e.g. other plants or crop waste – cellulosic ethanol) that might lead to increased efficiencies. It seems likely that the student is repeating information that he or she heard in class (*e.g.* environmental impacts, often drilled into students beginning with their first science class) as opposed to a dispassionate cost-benefit analysis. There is no evidence that the student is able to build an answer based on the conceptual foundations of the problem (conservation of mass, CO_2 fixation and release upon combustion, mechanisms and costs of generating ethanol versus gasoline). That the test graders considered this students' response one of the best suggests a misplaced value on reflex responses rather than on conceptual understanding. Therefore, it is no surprise that the student answered the way he or she did.

Rather than blame the student, we need to reexamine the question - what, exactly, are the key ideas a student needs to grasp to be able to analyze "the relative cost-benefits associated with using biologically-derived ethanol versus fossil fuels?" Based on our experiences, posing the question in a different manner may well have provoked a more conceptually informative response. As it is, our research suggests that a typical student would be disinclined to answer the ethanol question conceptually since they have already been told the context of the expected answer (*i.e.*, corn as a source of ethanol).

Creating Better Conceptual Questions.

Writing a test with questions that require students to use, and so reveal, their working understanding of a topic can be difficult and often involves multiple rounds of analyzing student responses and revising question. At the same time questions that elicit a reflex response can give the illusion of understanding, are easier to write, are easier for the student to answer, and easier for the instructor to grade. In our diffusion example, we soon realized that students understood that membranes and gradients were associated with diffusive events, but we learned little about whether students understood the root cause of diffusion: the constant motion of molecules. We needed to generate a question that would force students to answer *why* diffusion occurred. Clearly, simply asking "why?" did not work, so we asked a different type of question, a question based on a scenario that students were unlikely to have encountered previously - *"Imagine that you are molecule of ADP inside a cell. Describe how you manage to "find" an ATP synthase, so that you can become an ATP.*" The correct answer would be by random motions, or by diffusion. Surprisingly, we found that only approximately five percent of students mentioned the possibility of ADP finding the synthase by chance or diffusion. Most students described ADP as "looking" for the synthase, or hydrogen, concentration, or charge gradient. Thus, the answers to the new diffusion question revealed much more about students' understanding of the target concept than the answer to the first question.

Building Concept Inventory Questions

As we begin to understand how students approach a particular subject, we move on to develop one or more multiple choice questions as part of the BCI. In such questions, the incorrect answers or "distracters," are based upon our database of student responses. These distracters are truly distracting because they represent commonly held alternative responses to the question. Moreover, all parts of the question are written in student language that reduces the chances that students can recognize, rather than know, the correct answer. For example, in the final diffusion question (see below), both the correct response (e) and the distracters (a-d), were based on common alternative conceptions and presented in the students' own words. The final question becomes:

Imagine that you are molecule of ADP inside a cell. Describe how you manage to "find" an ATP synthase, so that you can become an ATP.

a. I follow the hydrogen ion flow
b. ATP synthase grabs me
c. My electronegativity attracts me to the ATP synthase
d. I would be actively pumped into the right area
e. By random motion or diffusion

We next tested this question for accuracy and validity through one to one and group interviews with students (think-alouds). During these interviews, we ask students what they think the question is asking, what they think each possible answer means, which answer they think is correct, and why they chose that answer. From interviews we found that students understand the meaning of the ATP question and the answer choices. However, the majority of students did not think that "e" was correct. Thus, our suspicion regarding most students' understanding of diffusion was confirmed: many students do not understand the fundamental concept behind diffusion.

Why Does Valid Conceptual Assessment Matter?

A cursory analysis of a number of science standards exams suggests that they are heavy on evaluating the recall of terms and facts and weak on the assessment of whether students have a working understanding of conceptual foundations. They tend to be reflexive in their approach, and as such measure whether a student has been paying attention to the material presented, rather than whether they understand how to use that information. Their structure appears to have a pernicious effect on our K-12 educational system. As political demands for higher student performance on these exams increases, teachers are both encouraged and coerced to teach to the test rather than to concept mastery. While it is possible to efficiently drill students on vocabulary (and reflexive responses), leading them to concept mastery is a much more time-consuming process. We would argue that many current testing instruments act to decrease the average students' competence, even as they increase their nominal achievement, because they encourage excessive (reflexive) content at the expense of conceptual confusion. In the absence of valid and confidently held understanding, students tend to generate scenarios that obscure rather than illuminate problem-solving, leading to uncertainty and frustration. If the goal of science education is for students to be able is to apply scientific understanding and analysis to new situations, it is critical that both instructors and evaluators become explicit in what tasks they expect their students to perform. Only then will curriculum and assessments positively reinforce one another, and provide the context for something better, and arguably more engaging, than memorization.

ACKNOWLEDGMENTS

We thank Isidoros Doxas, Valerie Otero, Steve Pollock, Noah Finkelstein and the LA-TEST/PER groups for helpful discussion and continuing insights. We thank Rebecca Klymkowsky for editorial advice. RG was supported by a Noyce fellowship and the BCI project is funded in part by a grant from the National Science Foundation.

REFERENCES

1. L. Shepard, *Educational Leadership* **46,** 4-9 (1989).
2. J.F. McClymer and L.Z. Knowles. J. *Excell. Coll. Teaching* **3**, 33-50 (1992).
3. Committee on Prospering in the Global Economy of the 21st Century, *Rising Above The Gathering Storm: Energizing and Employing America for a Brighter Economic Future*. (National Academies Press 2006).
4. D. Hestenes, M. Wells and G. Swackhamer, *The Physics Teacher* **30**, 141-166 (1992).
5. D. Hestenes and I. Halloun, *The Physics Teacher* **33**, 502-506 (1995).
6. E. Mazur, *Peer Instruction: A User's Manual*. (Prentice Hall, 1997).
7. R.R. Hake. *Am. J. Physics* **66**, 64-74 (1998).
8. D.L. Anderson, K.M. Fisher and G.J. Norman, *J. Res. Sci. Teaching* **39**, 952-978 (2002).