BIOCHEMISTRY AND MOLECULAR BIOLOGY EDUCATION
Vol. 36, No. 1, pp. 16–21, 2008

Articles

Student Assessment in Large-enrollment Biology Classes*

Received for publication, May 30, 2007, and in revised form, September 21, 2007

Manuel Alonso‡¶, Carlos Stella§, and Lydia Galagovsky‖

From the ‡Departamento de Ciencias Biológicas, Ciclo Básico Común, Universidad de Buenos Aires, Pab. III, P. B., Av. Int. Cantilo s/n, Ciudad Universitaria, 1428 Buenos Aires, Argentina, §Departamento de Bioquímica Humana, Facultad de Medicina, Universidad de Buenos Aires, Paraguay 2155, 5to. Piso, 1121 Buenos Aires, Argentina, and ‖Facultad de Ciencias Exactas y Naturales, Centro de Formación e Investigación en Enseñanza de las Ciencias (CEFIEC), Universidad de Buenos Aires, Pab. II, P. B., 1428 Buenos Aires, Argentina

Enrollments into first-year university biology courses may be very large, and therefore evaluating student learning can represent quite a challenge. In this article, we present our experience in assessing students by means of an assessment instrument called "Understand Before Choosing" (UBC). It has been used for six semesters, and its performance has been compared with two other common means of assessment, the use of multiple-choice questions and the use of open-ended questions. UBC consists of a text (100 lines, nearly 700 words) on the subject being tested, and a set of carefully worded questions that require the selection of one of five crafted options to be answered. To choose the best option, a student needs to understand the concepts embedded in the text.

Keywords: Assessment, cognitive skills, fast and objective assessment.

Assessment is a difficult task for teachers. Much energy and attention is currently being focused on improving the validity, reliability, and objectivity of the assessment tools.

In teaching science, one must focus on developing the students' capacity for analytical thinking, and comprehension of concepts presented, their ability to apply their knowledge and problem-solving skills [1–3]. Therefore, whenever an assessment tool is chosen, it should assess these objectives and not just facts and rote memory [4, 5].

There are different kinds of tools that are useful in evaluating students; for example, multiple choice, short answer or essay questions, practical examinations, questionnaires, peer rating, portfolios, interviews [1, 6]. Each of them has advantages and disadvantages. For instance, tests with multiple-choice questions are often used when a great number of students are to be evaluated; they are easy and fast to correct as well as objective [6].

Another type of traditional testing tool is the use of open-ended questions, for which students need to write responses of several lines. This type of device may involve rote learning as well as other levels of cognitive skills. However, it is difficult to mark objectively, either when this task is carried out by different teachers, or when the same teacher cannot maintain the same correction criteria between the first and the last tests.

These methodological considerations and the large number of students in our courses have led us to become interested in finding a quick and objective way of evaluating knowledge, comprehension, and ability to apply the material taught.

In this article, we present a method of assessment that we have called "Understand Before Choosing" (UBC),[1] in which we combine the benefits of the traditional multiple-choice test with those of the open-ended question test. We have used UBC to assess students in our crowded first-year biology courses at the University of Buenos Aires (1,000 students per semester) for over six semesters.

STRUCTURE OF THE UBC ASSESSMENT INSTRUMENT

UBC consists of a text of ~100 lines that describes a concrete example related to the topic being evaluated (metabolism of a certain cell type, the functioning of an organ or a system of organs). The test is divided into short paragraphs (15–20 lines, nearly 700 words) accompanied by a closely related set of questions that require the selection of one of five response options, as in more traditional multiple-choice test.

The text as a whole offers a specific case within a common context, which must be understood in its entirety. Students must relate their knowledge with the information provided in the text and with the question posed. From this association, they will derive the necessary clues that permit them to choose the correct

* This work was supported by grant UBA CyT 2004/07 X-181 from the University of Buenos Aires.

¶ To whom correspondence should be addressed. E-mail: manuelalonso64@yahoo.com.ar.

[1] The abbreviations used are: UBC, understand before choosing; DC_R, discrimination ratio; DF_R, difficulty ratio.

DOI 10.1002/bmb.20147

Box 1. Open ended questions involving an enzyme biochemistry subject.
Answer the following questions: a) What is/are the hydrolysis product/s of an enzyme? b) Explain the way by which an enzyme catalyzes a reaction. c) Which is the ATP function? d) How can you explain the specificity of an enzyme? e) Explain the activity regulation of an inducible enzyme. f) Define specificity of an enzyme. g) Explain why the reaction rate of an enzyme may not change although a great quantity of substrate is added. Justify.

answer. In this way, choosing the correct answer is dependent not only on the information retained in memory but also on the abilities acquired in the course to understand an explanation in scientific language and to use what has been learned to make further logical deductions.

COMPARATIVE DESIGN OF OPEN-ENDED QUESTION EXAMS, TYPICAL MULTIPLE-CHOICE QUESTIONS, AND UBC DEVICES

Analysis of a Concrete Example of UBC

To compare the design of open-ended question exams, multiple-choice tests, and UBC instruments, we present an example of three questions involving the same subject. Boxes 1–3 present questions in three formats. Each related to enzymatic biochemistry. Correct answers are highlighted in bold in Boxes 2 and 3.

These three devices have been sequentially used with freshmen in our biology courses. Open-ended question exams were used from 1996 to 2000, UBC between 2001 and 2003, and multiple-choice questions during 2004–2005. The main features associated with each type of exam are compared in Table I.

In Box 2, we present an example of a section of a UBC style test, illustrated here with a question on enzymatic biochemistry. The entire text consists of ~100 lines and refers to the biochemical functions and cellular characteristics of the liver. The topic of metabolism is presented in the course in the study of general processes, but neither metabolism itself nor the characteristics of the hepatocyte are specifically explained. On analyzing the example, the reader must keep in mind that this question was used with novice students in biology and biochemistry. This example was therefore challenging for them.

The text, which would read as a simple and "nice" introduction for a biochemistry student, is actually, in this context, rather difficult for a student who has previously only come into contact with biochemistry-related topics in a general way.

Analyzing our example, in the first question, the student must know that glucokinase is a protein, and there-

TABLE I

Comparative main features of open-ended questions, multiple-choice questions, and UBC instruments.

The open-ended questions are relatively easy to produce to assess different cognitive aptitudes. However, its correction is laborious and nonobjective. Multiple-choice questions and the UBC are easy to correct and objective but difficult to produce in a manner that enables them to assess competency in understanding and application of material learned. The UBC, however, provides the student with a specific case that integrates various concepts and which the student must solve by using his competency in information, comprehension, and application of appropriate concepts

Open-ended questions	Multiple-choice questions	UBC
Questions are easily formulated and may be not related to each other.	Questions are highly structured and may be not related to each other.	A small scientific text should be drawn up or selected. This text should not have been seen before by students. Questions are highly structured and require a comprehension of the text to be answered because they are related to each other.
Responses vary widely due to the different capacities of students in answering them.	Questions require a careful syntax to guarantee the same level of difficulty for each distractor to avoid skewed responses.	Questions require a careful syntax to guarantee the same level of difficulty for each distractor to avoid skewed responses.
Correction is not objective; there are difficulties among teachers concerning agreement on the final grading of each answer. There is a need to interpret different written expressions of the answers.	Objective and easy correction.	Objective and easy correction.
Enables evaluation of competency in communication, rote learning as well as different levels of cognitive skills such as comprehension, application, and content synthesis.	Questions may evaluate different levels of cognitive skills.	Questions point to different levels of cognitive skills within a common context, which places the student in a specific case.
Recommended for small classes.	Recommended for large classes.	Recommended for large classes.

BAMBED, Vol. 36, No. 1, pp. 16–21, 2008

Box 2. Understand Before Choosing (UBC) involving an enzyme biochemistry subject.

The text:

Hepatocyte metabolism, like cells of other tissues, requires energy supplies, for example, glucose. Once inside the cell, glucose is modified to avoid its exit, according to the following reaction:

Glucose + ATP → Glucose-6-P + ADP

This reaction is catalyzed by glucokinase. This molecule has a high Km and only phosphorylates glucose. The number of copies of glucokinase within the cell may increase or decrease in different physiological conditions.

In cells of other tissues, however, hexokinase catalyzes phosphorylation of glucose, mannose and fructose and has a Km lower than glucokinase.

After reading the text, answer the following questions:

1) Hydrolysis of glucokinase produces:
a) amino acids and ATP,
b) fatty acids,
c) amino acids,
d) monosaccharides,
e) nucleotides.

2) Glucokinase is:
a) a molecule that increases the activation energy of a reaction
b) a molecule that decreases free energy,
c) a molecule that decreases the activation energy of a reaction,
d) a cofactor that catalyzes a reaction,
e) a molecule that increases the free energy of a reaction.

3) ATP used in a phosphorylation reaction is:
a) long time energy storage,
b) coenzyme,
c) biological catalyst,
d) energy intermediary between catabolism and anabolism,
e) an amino acid.

4) According to the text above, the activity of glucokinase may be regulated by one of the following mechanisms:
a) activation by precursor,
b) negative feedback,
c) activation of zymogene,
d) gene control,
e) covalent modification.

5) Select the characteristic of glucokinase in relation with hexokinase
a) is more specific,
b) is less specific,
c) has a higher Vmax,
d) has more affinity for glucose,
e) requires less glucose concentration to reach Vmax/2.

6) If the velocity of glucokinase does not change although a great quantity of glucose and ATP is added, you can conclude that:
a) the enzyme increased its Km value,
b) the enzyme got denaturated,
c) the Km value of the enzyme has decreased,
d) the specificity of the enzyme has been modified,
e) all the active sites of the enzyme are bound to substrate.

fore its hydrolysis will produce amino acids. Because in many cases the students do not manage to associate and differentiate the function (enzyme) with the chemical nature (protein), this question allows us to assess if this association and differentiation have been successfully achieved.

The second question lets us evaluate whether the student has understood that glucokinase, as a catalyst,

Box 3. Multiple Choice Questions involving an enzyme biochemistry subject.
1) Hydrolysis of an enzyme produces: a) amino acids and ATP, b) fatty acids, **c) amino acids,** d) monosaccharides, e) nucleotides. 2) An enzyme is: a) a molecule that increases the activation energy of a reaction, b) a molecule that decreases free energy of a reaction, **c) a molecule that decreases the activation energy of a reaction,** d) a cofactor that catalyzes a reaction, e) a molecule that increases the free energy of a reaction. 3) ATP used in a phosphorylation reaction is: a) long time energy storage, b) coenzyme, c) biological catalyst, **d) energy intermediary between catabolism and anabolism,** e) an amino acid. 4) The activity of an inducible enzyme is regulated by one of the following mechanisms: a) activation by precursor, b) negative feed back, c) activation of zymogene, **d) gene control,** e) covalent modification. 5) The specificity of an enzyme means: **a) it binds to one kind of substrate,** b) it binds to different kind of substrates, c) it needs a small concentration of substrate to reach Vmax/2, d) it needs a great concentration of substrate to reach Vmax/2, e) it reaches Vmax/2. 6) If Vmax is constant: a) the enzyme increased its Km value, b) the enzyme got denaturalized, c) the specificity of the enzyme has been modified, d) the Km value of the enzyme has decreased, **e) all the active sites of the enzyme are bound to substrate.**

works by decreasing the activation energy of the reaction. Therefore, in this question, the idea is to determine if the student has successfully related catalysis with the function of the enzyme and then been able to indicate that what a catalyst does is what the enzyme does.

Although in the proposed reaction ATP is required, it neither catalyzes nor functions as a cofactor, which means that in question 3, the student must eliminate any option that does not coincide with what he has already learned: that ATP is an energy intermediary.

The following three questions require a closer reading and more extensive understanding of the text. The students have not studied the kinetic properties of either glucokinase or hexokinase; this is the first time they have been given a description of these enzymes. Therefore, question 4 can only be answered correctly if the student infers from the text that the increase or decrease of the levels of glucokinase, in accordance with physiological conditions, is due to a mechanism of genetic regulation. This question illustrates the many ways students can confuse issues. We have observed that in many cases the student confuses the reaction catalyzed by the enzyme (phosphorylation) with mechanisms that regulate enzyme activity and ends up answering that it is covalent modification (by phosphorylation). Other students state that the fact that the number of copies of the enzyme increases or decreases is related to the regulation by the final product, confusing the increase of the product of the reaction with the increase in the number of copies of the enzyme that catalyzes the reaction.

The fifth question requires a thorough analysis of the information given in the text. The student must distinguish between the specificity and the affinity of a substrate and apply the concepts of Km and enzymatic specificity.

Finally, question 6 requires students to have a clear understanding of which are the substrates of the enzyme (which they must deduce from the equation) and to relate these data with their theoretical knowledge. The answer can be deduced as follows: when the concentration of the substrates increases, the overall velocity does not change because all the active sites are bound to the substrate molecules. Again, in this case, the information provided is not sufficient on its own; the student's comprehension is also being evaluated.

DISCUSSION

Common multiple-choice questions consist of two parts: a statement or question and various response options. The statements or questions provide introductory information and may take the form of either a direct question or an incomplete statement. The response options contain one correct answer and several distractors. The options available are related to the context of the question or statement immediately preceding them, which guides the student with details pertaining to a specific situation. To answer a multiple-choice question well, the student must make the correct association between the "clue" given in the question and one of the options.

In the UBC, the questions or statements are related to a main text, which describes something specific (hepatocyte, erythrocyte, neuron, an organ, or particular organ system, a biological process). The information provided in the text can include structural, chemical, biochemical, physiological, genetic, molecular, or cellular aspects of the specific case, which acts as a common thread.

The text on the test orients and guides students through a specific situation, which they must analyze to answer the questions. Students must read and interpret this concrete scientific information with which they have never previously come into contact.

The example analyzed here, part of a greater context, allows us to show how it is possible to write questions that require the student to localize information in his long-term memory and relate it to the text and each question asked or statement made [7].

Generally, in this introductory biology course, the contents of the syllabus and textbooks used [8–10] deal with universal processes of molecular and cellular biology and biochemistry. For instance, the biological processes are explained in a general manner and all the possible cases (kinds of regulation, types of kinetics, the concepts of affinity, specificity, etc) are presented without going into great detail about what takes place in every specific cell type of each tissue, system, or organism. This method of teaching aims to cover the broadest possible spectrum of information on the topic.

Therefore, when given a UBC style test, the student cannot simply locate the clue in the question stem or in long-term memory [7] that will permit the choice of the correct answer.

For students to determine what the clue is, they must understand the different pieces of the puzzle, which make up the text and relate these back to the particular stem of each question. These cognitive abilities demand recognition of encoded information stored in their long-term memory and produce new information within the capacity of short-term memory or working memory [7].

In this way, UBC enables assessment of not only memorized information but also comprehension and application of taught material and, at the same time, lends itself to easy and objective correction. In an open-ended question test, it is possible to assess comprehension and application skills based on a particular text. However, it is difficult to correct this type of test in an objective way, assuring that the same criteria are maintained throughout the correction of all tests.

Comparative Statistical Validation

An average of 30–40% of students did not pass the exam when they were evaluated with open-ended questions. On the other hand, 45–50% of students did not pass the exam when traditional multiple-choice questions were used. UBC was used from 2001 to 2003 with 30–40% of students failing, a proportion similar to the open-ended questions test.

To evaluate the quality of this new device, discrimination and difficulty ratios have been calculated. Table II shows such ratios for the UBC question in Box 2. The whole exam consisted of 35 questions of the UBC type with five response options: only one correct answer and four distractors. Fifty-two random exams (of 600) were ordered from higher to lower total mark. Discrimination and difficulty ratios (DC_R and DF_R) were calculated involving the partial marks of the six questions shown in Box 2, considering the 27% highest and 27% lowest marked exams ($n = 14$), according to the following formula [11]:

$$DC_R = (R_H - R_L)/N_H$$

and $DF_R = (R_H - R_L)/N_{(H+L)}$, respectively, where DC_R is the discrimination rate, DF_R is the difficulty rate, R_H is the number of right answers of the highest group, R_L is the number of right answers of the lowest group, N_H is

TABLE II
Discrimination and difficulty ratios of the six questions analyzed in Box 2.
These ratios were calculated comparing the 27% highest and 27% lowest marked exams out of 52 ($n = 14$) according to the following formula [11]: $DC_R = (R_H - R_L)/N_H$ and $DF_R = (R_H) - R_L/N_{(H+L)}$, respectively, where DC_R is the discrimination rate, DF_R is the difficulty rate, R_H is the number of right answers of the highest group, R_L is the number of right answers of the lowest group, N_H is the number of individuals of any of the groups (14), and $N_{(H+L)}$ is the number of individuals of both the highest and the lowest groups

Question number	R_t	R_L	R_H	DC_R	DF_R
1	5	1	4	0.43	0.36
2	6	0	6	0.86	0.43
3	3	0	3	0.43	0.21
4	8	1	7	0.86	0.57
5	8	2	6	0.57	0.57
6	7	1	6	0.71	0.50

the number of individuals of any of the groups (14), and $N_{(H+L)}$ is the number of individuals of both the highest and the lowest groups.

Average discrimination ratio and average difficulty ratio on the whole exam were 0.44 and 0.58, respectively. Results show that values are in agreement with DC_R and DF_R standards [11].

FINAL REMARKS

Most teachers know how difficult it is to prepare a set of multiple-choice questions that demand other cognitive skills apart from recall information stored in long-term memory [7]. However, this type of exam generally does not relate all the questions to each other with a common thread. This is what we have proposed to do with the UBC, as it requires the students to constantly make use of both their comprehension of the text and previous knowledge to choose the correct option in answering each question.

Azer [12] and other authors such as Szeberényi [13] have emphasized the importance of problem-based learning and integrating evaluation as a part of the teaching–learning process. Teachers should set up scenarios that allow for the formulation of questions that test the integration and application of knowledge and text comprehension.

Students usually say that they prefer open-ended question exams because they can write down everything they have studied; multiple-choice questions, on the other hand, are more difficult for them to solve especially when they must sort through a lot of information stored in their long-term memory to select the best option (they are often similar in syntax and cause students to doubt). UBC showed similar standard marks as those achieved by students with open-ended question exams, which clearly show that it is possible to assess cognitive skills such as comprehension and application, relating all the material caught in a single test.

We have qualitatively analyzed students' opinions regarding our method of assessment. In general, they feel that it is more demanding than other methods we have discussed in this work, and they maintain that the UBC exam makes them think and reflect more on what they have learned. The students who were used to multiple-choice exams showed that they were trained to take tests that were based heavily on memorized material. With the UBC, however, they felt it was possible to relate or demonstrate what they had learned and at the same time be corrected in an objective manner.

As well, many students noted the degree of work that went into designing this type of test and took that as a sign of the teacher's dedication to assessing them in an objective way and in a manner that allows them to apply and relate all that they have learned during the course.

Previous studies by Azer [12] and Szeberényi [13] have suggested other examples of proposed testing devices that have been either attempted during only one semester or not tested at all. In this article, we have developed a systematic and sequential analysis over 3 years comparing UBC with open-ended question exams and traditional multiple-choice questions tests on the same subjects. In addition, this study has been done in massive classes demonstrating that UBC combines some virtues of open-ended question exams with the objective, quick, and easy correction of the usual multiple-choice questions instruments. The main difficulty of UBC is to draw up or select the proper text. One way to overcome this obstacle is to select material from scientific books or develop questions out of a combination of parts of the Experimental and Results sections of scientific articles.

We believe that presenting students with a specific situation in a context that they must analyze and understand allows for an improved assessment of what they have learned in terms of information, comprehension, and ability to apply acquired knowledge. The UBC enables this to be done in a quick and objective way and with a consistent criterion for correction from the first to last tests in courses with large numbers of students.

REFERENCES

[1] M. D. Sundberg (2002) Assessing student learning, *Cell Biol. Educ.* **1**, 11–15.

[2] E. Kitchen, J. D. Bell, S. Reeve, R. R. Sudweeks, W. S. Bradshaw (2003) Teaching cell biology in the large-enrollment classroom: Methods to promote analytical thinking and assessment of their effectiveness, *Cell Biol. Educ.* **2**, 180–194.

[3] S. A. Raizen, J. B. Baron, A. B. Champagne, E. Haertel, I. V. S. Mullis, J. Oakes (1991). Assessment in science education: The middle years. The National Center for Improving Science Education, The Network, Andover, MA.

[4] P. Black (1998) Assessment by teachers and the improvement of students' learning. In B. J. Fraser, K. G. Tobin, Eds. International Handbooks of Science Education, Part 2, Kluwer, London, pp. 811–822.

[5] J. Carneson, G. Delpierre, K. Masters (1996) Designing and Managing Multiple Choice Questions, Appendix C: MCQs and Bloom's Taxonomy (http://web.uct.ac.za/projects/cbe/mcqman/mcqman01.html).

[6] Saskatchewan Education (1991) In S. K. Regina. Student Evaluation: A Teacher Handbook: Saskatchewan Education, Canada.

[7] K. A. Ericsson, H. A. Simon (1999) Protocol analysis. Verbal Reports as Data, 3rd ed., MIT Press, Cambridge, Massachusetts.

[8] http://www.cbc.uba.ar/dat/cbc/con-min.html#Biologia.

[9] N. Campbell, J. B. Reece (2004) Biology, 7th ed, Benjamin Cummings, San Francisco, CA.

[10] W. K. Purves; D. Sadava, G. H. Dorians, H. C. Heller (2003) Life, The Science of Biology, 7th ed., Sinauer Associates and W. H. Freeman, New York, NY.

[11] N. E. Gronlund (1993) How to Make Achievement Tests and Assessments, Allyn & Bacon, Boston, MA.

[12] S. A. Azer (2003) Assessment in a Problem-Based Learning Course, *Biochem. Mol. Biol. Educ.* **31**, 428–434.

[13] J. Szeberényi (2003) The Role of a Chaperone in Protein Synthesis, *Biochem. Mol. Biol. Educ.* **31**.

REFERENCES USED IN CREATING UBC TEXT AND QUESTIONS IN BOX

T. M. Devlin (1999) Textbook of Biochemistry with Clinical Correlations, 4th ed., Wiley-Liss, New York.

D. Voet, J. G. Voet (1995) Biochemistry, 2nd ed., Wiley, New York.