Using Data Analysis Problems in a Large General Microbiology Course

Charles E. Deutch

A MAJOR goal of biological education should be to assist students in developing the ability to think critically (Moore 1993). They should be able to place new information in a broad scientific framework, use factual material they have learned in new contexts, interpret experimental data presented as graphs or tables, and draw appropriate conclusions from a series of observations. Student success in developing these skills is often evaluated on standardized tests through data analysis problems. A data analysis problem usually consists of an introductory paragraph describing a particular situation or experiment, one or more figures or tables of data, and a series of multiple-choice questions. Data analysis problems form a major part of the Science Test offered by the American College Testing (ACT) Program and the Scholastic Assessment Tests (SAT I and SAT II) offered by the Educational Testing Service. They also make up a large component of the Graduate Record Examination (GRE) and the Medical College Admissions Test (MCAT).

Although data analysis problems play an important role in the admission of students to undergraduate, graduate or professional programs, they are still not widely used in introductory biology courses. This is particularly true in the area of microbiology. A survey of the test banks and student guides that accompany the most commonly used textbooks in this field reveals an almost total lack of data analysis material. The questions in these test banks and student guides are primarily designed to assess mastery of basic vocabulary and factual material. While this is certainly important in an introductory course, the inclusion of data analysis problems is essential as well. These problems provide a mechanism for helping students move beyond rote memorization. They give them practice in critical thinking and illustrate the process of scientific inquiry. Accordingly, I have made an effort in the last several years to include data analysis problems in a large General Microbiology course. This article describes my experience with this course and presents several ways of incorporating data analysis problems in introductory courses.

Charles E. Deutch was Associate Professor in the Department of Biological Sciences at the University of Nevada, Las Vegas, Las Vegas, NV at the time this was written. His current address is the Department of Biology, Carleton College, Northfield, MN 55057.

Course Organization

General Microbiology at the University of Nevada, Las Vegas, is a one-semester, sophomore-level course with prerequisites of one semester of general biology and one semester of college chemistry. Three lecture sections are offered each year by different faculty, with enrollments of 50–150 students per section. Students from the various lecture sections register separately for a common laboratory component. About half of the students enrolled in the course are biology majors; the others are majors in allied health areas such as nursing and clinical laboratory sciences. The course is a general survey of microbiology which covers the structure of prokaryotic cells; the growth, metabolism and genetics of microorganisms; the biology of viruses; the interactions of microbes with plants or animals (including their pathogenicity); and the uses of microorganisms by humans. Four examinations covering the lecture material are usually given during the semester: three mid-term exams and a two-hour final exam. Because of the large enrollments in the course, the exams are often machine graded.

Data analysis problems have been used in my sections of the course in two ways. In 1992, these problems were simply included as a component of each examination. Most of each exam (80%) consisted of "standard" multiple-choice questions designed to test mastery of essential vocabulary and understanding of basic concepts. The remaining 20% consisted of two data analysis problems, each containing five multiple-choice questions. The answers to the data analysis problems were posted along with the answers to the other exam questions, but they were not discussed further. In 1994, the next time I taught the course, the data analysis problems from 1992 were included in the course packet distributed to students through the bookstore. Students were asked to work through each of the problems that complemented the more descriptive lecture material as a homework assignment. Each problem was then discussed at length in class. The examinations consisted of standard multiple-choice questions (80%) and questions based on new data analysis problems (20%). All answers were posted, but answers to the data analysis problems were also discussed in detail when the exams were returned.

Structure of Data Analysis Problems

Each data analysis problem is usually based on a single journal article. Although recent articles are often used to give students a sense of what is going on in microbiology today, older papers are sometimes included that better illustrate the lecture material. Articles from almost any scientific journal can be used. However, because each data analysis problem needs to be self-contained and accessible to the students, it is often easier to use short articles or notes rather than full-length papers. It is also sometimes necessary to edit the data to keep the problem a reasonable length.

Three types of questions are incorporated into each data analysis problem. Type A questions are those that ask students to recall basic terms or concepts which are relevant to the problem. This type of question is designed to determine if students can use information from the course in a new context. Type B questions are those that ask students to read a table or figure of data. This type of question is designed to determine if students can see relationships presented in a tabular or graphical form. Type C questions ask students to draw conclusions from the data presented, to make predictions from the results shown, or to integrate the material with other information presented in the course. This type of question is designed to determine if students can think analytically or critically.

Three sample data analysis problems are shown in the Appendix as Examples 1, 2 and 3. The first problem deals with the structure and function of bacterial capsules and includes a relatively simple data set. Questions 1 and 2 are Type A; Question 3 is Type B; and Questions 4 and 5 are Type C. The second problem deals with bacterial fermentation and includes a more complex data set. Questions 1 and 3 are Type A; Questions 2 and 4 are Type B; and Question 5 is Type C. The third problem deals with microbial pathogenicity and contains several data sets that are presented sequentially. Question 1 is Type A; Questions 2, 3 and 4 are Type B; and Question 5 is Type C. While designating each question as Type A, B or C is sometimes a little arbitrary, it is helpful to plan on testing critical thinking skills in several different areas. The three examples illustrate the range in difficulty which is possible in constructing these problems. The problems used early in the semester tend to be relatively easy; they then increase in difficulty as the course goes along.

Student Performance on Data Analysis Problems

The performance of the students on the exams given in 1992 and 1994 is summarized in Table 1. The standard deviations of the means were very large for both the standard questions and the data analysis problems, reflecting both differences in student ability and variations in difficulty. Nevertheless, several trends were apparent. The percentage of students answering the standard multiple-choice questions correctly on each exam was about the same each year. However, the students did noticibly better on the data analysis problems in 1994 than in 1992. In 1992, the percentage of students answering the data analysis questions correctly was lower than the percentage for the standard questions. In 1994, the percentage of students answering the data analysis questions correctly was about the same as that for the standard questions, except for the second exam which had two rather difficult problems. Table 2 shows a breakdown of student scores by the type of question on the data analysis problems. Performance on Type A and Type B questions was similar, but scores on Type C questions tended to be lower. Again, however, students did much better in 1994 than in 1992.

The difference in student performance in 1992 and 1994 cannot be attributed to characteristics of a particular class. Performance on the standard multiple-choice questions was essentially the same in both years. The size of the class (about 150 students), the textbook, the lecture notes, and the class handouts were also the same in 1992 and 1994. The difference in performance appeared to be due primarily to test anxiety. Because the data analysis problems were a new element of the course in 1992, students were simply overwhelmed by them. They missed more Type A questions than expected, even though they had the information available. Type C questions were particularly difficult because students were not familiar with this type of analytical thinking. In 1994, a

Table 1. Overall summary of student test scores.

		Percentage Correct	
Year	*Exam*	*Standard Questions*	*Data Analysis Questions*
1992	1	74.8	67.4
	2	73.0	66.6
	3	64.7	54.4
	4	70.9	55.1
	mean	70.8	60.9
1994	1	74.5	71.1
	2	71.0	56.3
	3	68.8	69.6
	4	70.7	69.8
	mean	71.2	66.7

Table 2. Student test scores on data analysis problems by question type.

	Percentage Correct	
Question Type	*1992*	*1994*
A	63.9	71.0
B	68.6	71.1
C	52.4	62.2

much greater effort was made to prepare students for the data analysis problems. While they could not predict the subject areas or journals from which the data analysis problems would be drawn, they did know that such problems would be included on the tests. By assigning the old problems as homework and discussing them in class, the students learned how to recognize the types of questions that might be asked. Going through the data analysis problems after each exam also helped them prepare for the next exam. As in most other exam situations, practice can alleviate a great deal of anxiety.

Student Responses to Data Analysis Problems

Students have commented about the data analysis problems on the course evaluations distributed at the end of each semester. Even though the problems made up only 20% of each exam, students tended to see them as a larger factor in their final grade than they actually were. One student said, "The short passage part was too hard" and another said, "The tests were too long." Some students felt the data analysis problems were unfair because they were based on material not covered directly in class or in the textbook. However, another student said, "The tests were fair and made you think—I don't like those types of questions but they are a great learning tool."

Students also have commented about the data analysis problems more informally. Many of the biology majors in the course hope to go on to medical school or graduate school. After taking the MCAT or GRE tests, several students indicated they now understand why data analysis problems were included in the course. They said they wished that they could have had more practice in dealing with these problems and hoped more instructors would use this approach.

Advantages of Using Data Analysis Problems

The inclusion of data analysis problems in General Microbiology has three major advantages for an instructor. First, it provides a mechanism for presenting case studies that illustrate current issues in the field. In addition to the data analysis problems shown in this article, there have been problems on the uptake of uronic acids by the plant pathogen *Erwinia carotovora*, the adaptation of *Listeria monocytogenes* to cold temperatures, the metabolism of toxic heavy metals by photosynthetic bacteria, the quantitation of endotoxin activity, and the attachment of *Vibrio cholerae* to host cells. Second, the use of the data analysis problems is a way to discriminate between those students who understand the material and those who simply have memorized it. While I have no quantitative data to prove it, students who receive grades of "A" in the course usually do better on the data analysis problems than those who receive grades of "B" or "C." Finally, the inclusion of data analysis problems changes the dynamics of an introductory course to some extent. When students come for help during office hours, it is now sometimes because they want to discuss the problems further. While some students still ask, "What did you say on Wednesday after you said . . . " or "What is going to be on the next test," others now want to discuss the design of the experiments or other conclusions that might be drawn from the data. This has made teaching the class much more rewarding.

Alternative Approaches to Using Data Analysis Problems

The use of data analysis problems can be varied depending on the interests of the instructor and the size of the class. Instructors with a strong interest in pedagogy may want to use Bloom's (1956) taxonomy of educational objectives to devise questions that test student achievement at more defined levels. Instructors in courses with smaller enrollments may wish to ask the students to write out answers to the data analysis questions, in which they argue for or against each of the possible responses. They may also want to use the problems as the basis for small-group discussions, either in an open-ended format or as a series of individual student presentations. With small-enrollment classes, it would also be feasible to ask the students to write essays on the problems either as homework or as exam questions.

While this article deals specifically with the use of data analysis problems in a large General Microbiology course, this approach can be adapted to any course from General Biology to Plant or Animal Physiology to Ecology. The only difference would be the journals from which the problems are derived and the format in which the problems are presented. Pechenik and Tashiro (1992) recently described a similar approach for teaching problem-solving skills to small groups of students.

Conclusions

Data analysis problems can be used successfully in a large introductory course, even one in which the exams consist entirely of multiple-choice questions and out-of-class contact with the instructor is limited. Two factors seem to be important in ensuring that students have a reasonable chance of doing well on these problems. First, the data analysis problems need to be written carefully and to include data sets with which the students can deal. Sometimes I have used problems that, in retrospect, were probably beyond the grasp of most students. The problems involved techniques with which some students were unfamiliar or required peripheral information about which they were unaware. At the introductory level, the problems work best when they are simple and straightforward. Second, students need to be told from the beginning that data analysis is an important component of the course. Using old data analysis problems as homework assignments is a simple way to familiarize students with them. Ideally, there should be at least one problem for each lecture topic. Although discussing old problems or exam questions in class is also helpful, it does take time away from the lecture material. In 1994, students often did not attend the class in which the exam problems were discussed. In the future, more descriptive explanations about the homework assignments and exam questions will simply be posted and the amount of in-class discussion will be reduced.

Data analysis has long been a key element in courses in genetics, and books of problems have been written to accompany advanced textbooks in biochemistry and molecular cell biology (Freifelder 1987; Gumport et al. 1990; Wilson & Hunt 1994). Recently, a section on problem-based learning was added to the journal *Biochemical Education* (Smith 1993). Now is the time to extend this approach throughout the biology curriculum.

Acknowledgments

I thank an anonymous reviewer for suggesting the use of Bloom's taxonomy and for describing the incorporation of data analysis problems in courses at their institution.

References

Bloom, B. (1956). *Taxonomy of Educational Objectives, Handbook 1: Cognitive Domain.* New York: McKay.

Freifelder, D. (1987). *Study Guide for Molecular Biology 2/e.* Boston: Jones and Bartlett.

Gumport, R.I., Jonas, A., Mintel, R., Rhodes, C. & Koeppe, R.E. (1995). *Student's Companion to Stryer's Biochemistry,* 4th ed. New York: W. H. Freeman.

Moore, R. (1993). Doing more than just "covering" information. *The American Biology Teacher, 55,* 260–261.

Pechenik, J.A. & Tashiro, J.S. (1992). The graphing detective: An exercise in critical reading, experimental design & data analysis. *The American Biology Teacher, 54,* 432–435.

Smith, C.A. (1993). Features section: Problem-based learning. *Biochemical Education, 21,* 32–33.

Wilson, J. & Hunt, T. (1994). *Molecular Biology of the Cell: The Problems Book* (revised edition). New York: Garland Publishing, Inc.

Author's Note

In 1966, I taught the General Microbiology course again in the same basic format, although a different textbook was used and the course enrollment was only about 45 students. Data analysis problems from previous years were distributed and 20% of each exam again was based on new data analysis problems. The answers to the homework assignments and exam problems were posted, but the questions were not discussed in class unless the students specifically asked about them. The mean score on each exam again was about 70% and the mean for the data analysis problems was only slightly lower (69%). This suggests that data analysis problems can be used effectively without severely reducing the amount of class time available for lectures or other activities.

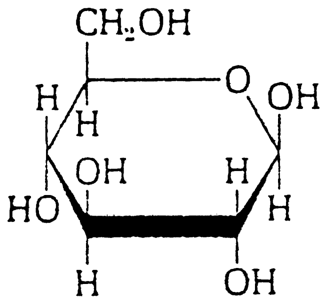
Appendix

Example 1. Data analysis problem in microbial structure.

Many microorganisms produce extracellular polysaccharides that form a coat lying outside of the cell wall. These polysaccharides are the primary point of contact with the external environment and so may be particularly important in cell-cell interactions. Cowpea rhizobia are a group of slow-growing, Gram-negative bacteria that can infect and form root nodules in a wide range of leguminous plants. These bacteria are often found in tropical soils and usually form wet slimy colonies on agar plates in the laboratory. Hollingsworth, Smith and Ahmad (*Archives of Microbiology, 142,* 18–20, 1985) compared the extracellular polysaccharides formed by six cowpea isolates. They measured the amount of polysaccharide produced by each organism and determined its chemical composition. The polysaccharides contained three types of monosaccharides: glucose, galactose and mannose. The results for three of the strains are shown below.

		% of Total Monosaccharides		
Strain	*Total Extracellular Polysaccharide (μg/ml)*	*Glucose*	*Galactose*	*Mannose*
A	350.0	48.5	27.2	24.3
B	26.8	—	—	100.0
C	87.5	—	38.1	61.9

1. The most common name for the polysaccharide coat formed by these bacteria is the:
   a. S-layer
   b. cell wall
   c. capsule
   d. cell membrane
2. The following figure shows the structure of glucose, a monosaccharide found in Strain A.



If glucose molecules were linked together to form a polysaccharide, what type of linkage would occur between them?
   a. ester linkage
   b. ether linkage
   c. peptide linkage
   d. amide linkage
3. Based on the data given in the table, which of the bacteria produces the most polysaccharide?
   a. A
   b. B
   c. C
   d. All of the bacteria are the same.
4. Suppose you found that Strain A could bind to the roots of particular plant species and induce nodule formation but Strains B and C could not. Which of the following hypotheses would be the most reasonable?
   a. The plant roots have receptors that specifically interact with glucose.
   b. The plant roots have receptors that specifically interact with galactose residues.
   c. The plant roots have receptors that specifically interact with mannose.
   d. The plant has no receptors on the root surface.
5. Suppose you now test this hypothesis by adding samples of Strain A to a series of pots containing the plant of interest. In some cases, you simultaneously add a solution containing individual monosaccharides. After two weeks, you obtain the following results.

Monosaccharide Added	*Percent of Plants with Nodules*
none	95%
glucose	10%
galactose	93%
mannose	96%

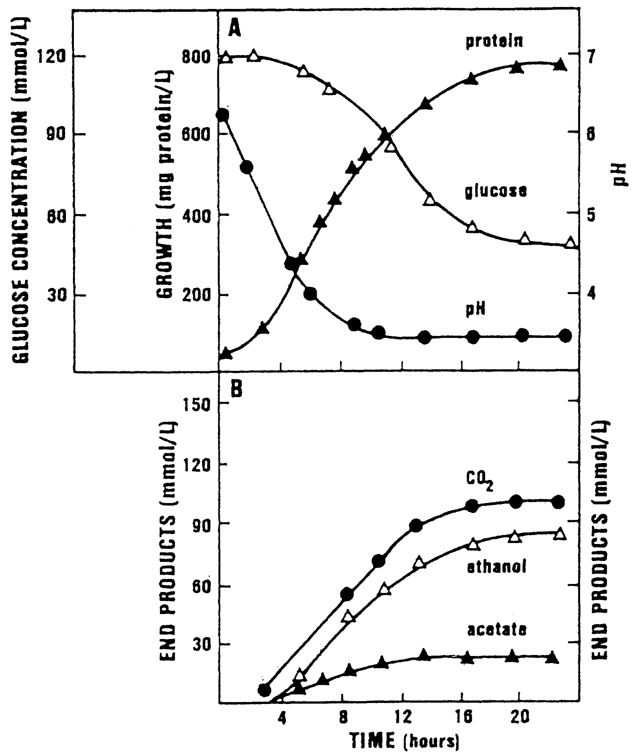
   a. These results support the hypothesis that the plant has receptors that specifically interact with glucose.
   b. These results support the hypothesis that the plant has receptors which specifically interact with galactose.
   c. These results support the hypothesis that the plant has receptors which specifically interact with mannose.
   d. These results support the hypothesis that the plant has no receptors on the root surface.

Answers:

1. c 2. b 3. a 4. a 5. a

Example 2. Data analysis problem in microbial metabolism.

Sarcina ventriculi is a Gram-positive bacterium that can utilize simple carbon sources such as glucose under anaerobic conditions. It is capable of growth at a wide range of pHs, from 2.0 to 10.0. Goodwin and Zeikus (*Journal of Bacteriology, 169,* 2150–2157, 1987) studied the growth and metabolism of this organism with the goal of understanding how it functions at low pHs. They inoculated 20 ml of an unbuffered glucose-containing medium with *S. ventriculi* and incubated the cells at 37° C in the absence of oxygen. They periodically removed samples from the culture for analysis, with the results shown in the following figure.



1. What type of culture was used in these experiments?
   a. a batch culture
   b. a synchronous culture
   c. a continuous culture
   d. a stationary culture
2. Which of the curves shown in the figure represents the growth of the microorganisms?
   a. glucose concentration
   b. protein concentration
   c. CO_2 concentration
   d. acetate concentration
3. From the introductory paragraph and the data given in the figure, what type of metabolism do these bacteria carry out?
   a. fermentation
   b. aerobic respiration
   c. anaerobic respiration
   d. photosynthesis
4. Based on these data, which of the following factors appeared to cause the bacteria to enter stationary phase?
   a. exhaustion of glucose
   b. accumulation of acetate (low pH)
   c. oxygen limitation
   d. none of the above
5. Goodwin and Zeikus concluded from these data that the metabolism of glucose occurred in two stages. Which of the following observations best supports this conclusion?
   a. Growth of these bacteria was slower at the beginning and eventually reached a maximal level.
   b. CO_2, acetate, and ethanol were all formed as the bacteria grew.
   c. Acetate formation reached a maximum at 10 hours, when the pH leveled off at 3.3, but growth continued up to 20 hours.
   d. Glucose was still available in the culture after 20 hours.

Answers:
1. a 2. b 3. a 4. d 5. c

Example 3. Data analysis problem in microbial pathogenesis.

Porphyromonas gingivalis is commonly found in the gum area below the teeth and has been regarded as a major cause of adult periodontal disease. This nonmotile anaerobic Gram-negative rod requires heme for growth and has been shown to degrade hemoglobin released from red blood cells. Hoshi et al. (FEMS *Microbiology Letters 114*, 273–278, 1993) recently showed that *P. gingivalis* produces toxin capable of causing hemolysis—the disruption of red blood cells.

1. What structure of the red blood cells must be disrupted in order to release hemoglobin into solution?
   a. cell wall
   b. cell membrane
   c. nucleus
   d. endoplasmic reticulum
2. To detect hemolysis, Hoshi et al. grew the bacteria in two different liquid media (designated Medium A and Medium B) and then mixed the cultures of *P. gingivalis* with red blood cells in varying combinations in a buffered saline solution. After 60 minutes, they added 3 ml of the saline solution to each mixture and centrifuged the samples for 5 minutes. They then measured the absorbance of the supernatant liquid at 541 nm to determine the amount of hemoglobin released. The results are shown below.

Conditions	*Hemolytic Activity*
red blood cells	0.05
culture grown in Medium A	0.00
culture grown in Medium A + red blood cells	0.16
culture grown in Medium B	0.00
culture grown in Medium B + red blood cells	1.62

From these results, one can conclude that:
   a. Most of the hemolysis is due to spontaneous breakdown of the red blood cells.
   b. *P. gingivalis* itself contains hemoglobin and lyses in the buffered saline solution.
   c. Cultures of *P. gingivalis* grown in Medium A have more lytic activity than cultures grown in Medium B.
   d. Cultures of *P. gingivalis* grown in Medium B have more lytic activity than cultures grown in Medium A.

3. To determine the location of the toxic material made by *P. gingivalis*, Hoshi et al. centrifuged the bacterial culture grown in Medium B and saved the medium for analysis. They then disrupted the bacteria with high energy sound waves, and centrifuged the mixture to separate the cytoplasmic contents from the broken cell envelopes. Each sample was analyzed for hemolytic activity, with the following results:

Sample	*Hemolytic Activity*
culture medium	1.58
cytoplasmic contents	0.04
broken cell envelopes	0.00

These results indicate the hemolytic toxin is located:
   a. in the culture medium
   b. in the cytoplasm
   c. in the cell envelope
   d. in the extracellular space

4. To determine the chemical nature of the toxin, the sample containing the toxin was treated with heat, a proteolytic enzyme called trypsin, a phospholipase, or an enzyme called amylase that degrades polysaccharides. The results are shown in the next table.

Treatment	*Relative Hemolytic Activity*
none	100%
heat (100° C, 10 min)	2%
heat (60° C, 30 min)	6%
trypsin (4 mg/ml)	0%
trypsin (0.4 mg/ml)	21%
phospholipase (4 mg/ml)	98%
amylase (4 mg/ml)	99%

Which of the following statements best describes the effect of heating on the toxin?
   a. The heat had no effect.
   b. The heat probably broke the toxin into its individual atoms.
   c. The heat probably broke the toxin into its component residues.
   d. The heat probably disrupted the three-dimensional structure of the toxin.

5. Based on the results shown in the table for Question 4, one can conclude that the toxin is probably a:
   a. protein
   b. phospholipid
   c. polysaccharide
   d. nucleic acid

Answers:
1. b   2. d   3. a   4. d   5. a