JOURNAL OF RESEARCH IN SCIENCE TEACHING VOL. 37, NO. 8, PP. 854–870 (2000)

Can Undergraduate Biology Students Learn to Ask Higher Level Questions?

Gili Marbach-Ad and Phillip G. Sokolove

Department of Biological Sciences, University of Maryland Baltimore County, 1000 Hilltop Circle, Baltimore, Maryland 21250

Received 14 June 1999; accepted 22 May 2000

Abstract: Our goals in this study were to explore the type of written questions students ask after reading one or more chapters from their textbook, and to investigate the ability of students to improve their questions during the course of a single semester. In order to classify student's questions we used a taxonomy that we have developed specifically for this purpose. Two comparable populations were examined: Undergraduate students in a large, introductory biology class who were taught in traditional lecture format, and students in a similar class who were taught in cooperative/active learning style. After the taxonomy was presented to the active learning class, more students were able to pose better, written questions. Their questions became more insightful, thoughtful, and content-related, and were not easily answered by consulting the textbook or another readily available source. The best questions could be recast as scientific research questions (i.e., hypotheses). In contrast, when the taxonomy was presented to students in the traditionally taught class, the quality of student-posed questions was largely unchanged. Various explanations for the difference in outcomes are discussed, and methods are suggested about how generally to encourage students' questions and to improve their question-asking skills regardless of overall teaching style. © 2000 John Wiley & Sons, Inc. J Res Sci Teach 37: 854–870, 2000

Introduction

"...And after reading this chapter give me your best question..." This was a homework exercise given to students in an introductory course in general biology. But what is meant by the term "best" question? Or even simply a "good" question? Can we define for our students what we consider to be good questions? Are there ways to teach or to encourage students to ask better questions? And, why should we want to encourage students to ask questions at all?

In recent years educators have emphasized and investigated the importance of students' questions in the teaching/learning process: in elementary schools (Commeyras, 1995), at junior high schools (Houston, 1938; Watts, Gould, & Aslop, 1997), and in high schools (Dillon, 1988; Dori & Herscovitz, 1999). A few studies have also been conducted at the college level (Pearson & West, 1991; West & Pearson, 1994).

A number of reasons have been expressed for the need to encourage student questions. First and perhaps foremost in science classes, is that emphasis on students' questions conveys the message that the science disciplines are areas where inquiry is a natural component and questions need constantly to be raised (Woodward, 1992 p. 152). As Orr (1999) points out, ‘‘Good science demands two things: that you ask the right questions and that you get the right answers. Although science education focuses almost exclusively on the second task, a good case can be made that the first is both the harder and the more important’’ (p. 343).

Second, in posing questions pupils in non-science as well as science disciplines both shape and expose their thoughts. Hence, student questions provide opportunities for teachers’ insight into students’ thinking and conceptual understanding (Woodward, 1992, p. 146). Third, our own thinking can be provoked by student questions to help us as teachers reach a broader understanding of text (Commeyras, 1995). Fourth, student-generated questions can increase their understanding and retention of textual narrative (Bean, 1985).

Although we would like students to learn science by asking appropriate questions, they are reluctant to do so. Dillon (1988) suggests that younger children ask questions, and that they ask more questions as they get older. They may indeed be asking questions as they read and study a textbook, or they may be asking questions of their friends and family, but they do not ask questions in the classroom. Studies show that ‘‘as grade level increases, students ask fewer on-task attention questions in class’’ (Good, Slavings, Harel, & Emerson, 1987, p. 186). Good et al. suggest this probably occurs because students do not want to call attention to themselves. It also seems that fewer questions are asked because teachers often ‘‘do not like’’ students to ask questions. Wood and Wood (1988) have speculated about this and other possible reasons that teachers do not encourage pupils to air their own views. But whatever the reason, the outcome is predictable: students learn not to ask questions in class.

Today, science education research tends to focus more on students than teachers, and the emphasis is on the learner and the learning process. From a constructivist perspective, learning outcomes are assumed to depend little, if at all, on what the teacher ‘‘presents’’ in lecture format. Rather, they are an interactive result of what information is encountered (based on guidance from the instructor) and how the student processes it (see Brooks & Brooks, 1993; Caprio, 1993; Johnson, Johnson, & Smith, 1991; Resnick, 1987; Yager, 1991).

Over the past several years, there has been a trend in science disciplines to move away from the more traditional, lecture method of teaching toward a more active approach (Cooper, 1995; Ebert-May, Brewer, & Allred, 1997). In a few designed studies that illustrate new approaches to teaching science, one finds an emphasis on the importance of the student’s questions. Most such studies have been conducted in laboratory classes (Dori & Herscovitz, 1999; Heady, 1993). They suggest that an effective strategy for improving the problem-solving ability is to foster students’ question-posing capabilities.

Encouraging students to ask questions and to seek information in order to answer their questions is certainly one way teachers can promote the notion of independence in learning. In our study we have tried to obtain answers to the following: (a) How can we define for introductory biology students what we consider to be a ‘‘good’’ question? (b) Are there ways to teach or encourage students to ask better questions?

In order to encourage desirable questions from students we first had to be able to distinguish the different categories of questions they might ask. This forced us to build a semi-hierarchical taxonomy, which in turn has allowed us to classify students’ questions. Bloom’s taxonomy (Bloom, 1984) and other methods that have been used to classify questions proved inadequate for our purposes because those classification schemes are designed (or at least have been used) mainly to categorize teachers’ questions rather than students’ questions. Methods proposed specifically to classify students’ questions were helpful to us in a general sense, but were found not to be appropriate for a college-level science class (Olson, Duffy, & Mack, 1985; Watts, et al., 1997). Composition-oriented guide books (i.e., Zinsser, 1976) that provide instruction about how to write in different disciplines such as travel, sports, or science deal primarily with descriptive presentation rather than with question-formulation. Pechenik (1993) recognizes that focusing on "the right question" is a key to writing a good research proposal, but provides no advice for the reader about how to differentiate among questions, or go about formulating "a valid and logically developed question (p. 143)."

In the study described here, which is part of a larger, longitudinal research project, we examined student questions in classes that expose freshman biology students to two different styles of teaching: traditional lecture and active/cooperative learning.

The traditional lecture style of teaching has been described as a format in which students are placed in a passive role that is not often conducive to learning (Cooper, 1995), but it is, without doubt, an efficient and cost-effective method for information delivery. Active learning using cooperative groups is an instructional technique that requires students to work together in small, fixed groups on a structured learning task (Cooper, Prescott, Cook, Smith, & Mueck, 1990). The value of cooperative learning has been highlighted by many researchers (e.g., Caprio, 1993; Johnson & Johnson, 1994; Johnson, Johnson, & Smith, 1991; Slavin, 1983). According to Cooper (1995) the advantages of cooperative learning are: students take responsibility for their own learning and become actively involved in the process; they develop higher-level thinking skills; cooperative learning helps to increase their retention; cooperative learning helps to increase student satisfaction; cooperative learning promotes positive attitudes toward the subject matter.

Johnson and Johnson (1994) have written that "cooperative learning has been around a long time [and] it will probably never go away (p. 38)." However, implementing cooperative/active learning methods is difficult in large science classes and few college instructors have taken on the challenge (see Cooper, 1995 and Ebert-May, Brewer, & Allred, 1997 for successful examples). Student interaction with the instructor or with each other may be difficult or impossible due simply to class size. Whole class participation is typically limited to a few, assertive individuals while the rest of the class plays a passive role. It is therefore particularly important to search for teaching methods that will counteract student passivity and place the student in a more active learning environment (Cooper, 1995).

In 1992 Kagan and Roberson wrote that active learning using cooperative groups was the most extensively researched educational innovation to date. At that time, there had been over 600 research studies of cooperative learning, and over half of those studies examined academic achievement. Here, in contrast, we have not studied academic achievement, but have instead examined how a specific intervention influenced question-asking in two different teaching environments and whether it could improve question-posing skills over a single semester.

Overview of Present Study

The present study was conducted in two classes: (a) a large, "active learning" introductory biology class offered in Fall 1998, and (b) a large introductory biology class offered in Spring 1999 taught in traditional, lecture format. Both classes were of similar size (225–250 students), were taught by the same instructor (Sokolove), and used the same textbook (Starr and Taggart, 1998).

Students in both active learning and traditional classes were told about the differences between the classes in the first or second class session. Every student in both classes was asked to sign a consent form that (a) informed them of the comparative research study in which we were engaged, (b) explained that they were being asked, but not required, to participate by completing (anonymous) surveys and allowing themselves to be interviewed, and (c) reassured them that the

Table 1
Comparison between the "active learning" class and the "traditional" class

Active learning class	Traditional class
• Students were challenged to ask questions in class about topics that were unclear or about issues that related one topic to another. Wireless microphones were used to enhance in-class communication.	• Few students asked questions in the classroom; other students could not hear them well and rarely responded with additional questions or clarifying statements.
• The instructor received many e-mail questions from students about the subjects that he taught.	• E-mail questions were almost exclusively procedural ("what chapters will be on the next exam?").
• In addition to chapter homework exercises students were given take-home assignments in which they were instructed to contribute two, original, written questions after they had completed a short research paper on a given subject (using as sources textbooks, the internet, journals, professional experts, newspapers, etc.). Assignments including the two questions were graded and handed back to the students.	• The chapter homework exercises that we describe in this report were the only written homework that the students were asked to do during the semester.
• Significant portions of some class sessions were devoted to helping students learn to recognize the difference between questions.	• When students handed in their questions there was no in-class discussion about them aside from a few comments that the instructor made about questions with misunderstandings or misconceptions.
• Students worked in cooperative learning (CL) groups (3–5 students) during class. They also attended one hour/week discussion sections of ~25 students, in which CL groupings were randomly formed at the semester's start.	• Cooperative learning groups were not used in the large class, but were employed by the teaching assistants in the smaller, weekly discussion sections.
• Students were accustomed to multiple types of assessment. In addition to exams and quizzes (35%) their final grade included scores on take-home assignments (30%), and on their participation in class (35%). Almost all students made an effort to complete all assignments and homework exercises.	• Final grades in class depended only on exams, and even when the instructor offered extra credit for formulating questions, only half of the students formulated questions.

results of our research would be used to improve teaching, but would in no way affect their grade for the course. They were also informed about the availability of the other class and were encouraged to enroll in the class that they felt matched their learning style. Institutional Research Board approval was sought and received for Human Subjects Research under a provision that permits broad discretion on gathering data for purposes of research on the educational effectiveness of new or innovative curricula and teaching methods.

The overarching issue in this active research study was whether student questions showed improvement when students were given information about the kinds of questions desired in their science course. The design of the study derives from the observation that in-class presentation of a student-question taxonomy (see below) in the Fall 1998 active learning class was followed almost immediately by improved written questions from students. The active learning class employs many innovative approaches including name badges (so students, TAs, and the instructor can get to know each others' names), wireless microphones (so that students can hear each other's questions and comments), and a class session about recognizing good questions (conducted early in the semester to begin to focus student attention on questions—both their own and those of others). Because in the Fall active learning class there are many innovations that emphasize question-asking, presentation of the classification scheme alone might or might not have been key element in fostering better student questions. Therefore, we tested the following hypothesis: Presentation of the classification scheme is sufficient to produce better student questions.

The test was conducted in the following semester's traditional class (Spring 1999) where the textbook and the instructor were identical, and the student population was equivalent (different students, but similar distribution by gender, race/ethnicity, SAT scores, etc.) The active-learning and traditional classes are part of a longitudinal comparative study that will examine long-term outcomes for student, in the two types of learning environments (see Table 1 for a summary of the differences). For purposes of the study reported here, the traditional class environment was ideal, since it purposely lacked the innovative approaches found in the active-learning class. Thus, the effect of presenting the classification scheme could be observed in the absence of confounding factors.

In order to classify student's questions we used a taxonomy that was developed for this purpose (Marbach-Ad & Sokolove, in press). The taxonomy includes eight categories of student questions:

Category 0: Questions that do not make logical or grammatical sense, or are based on a basic misunderstanding or misconception, or do not fit in any other category. (This is a "catch all" category that instructors can readily subdivide for teaching purposes—for example, when grading written questions. In this case we chose *not* to subdivide the category in order to focus on the characteristics of desirable questions.)

Category 1a: Questions about a simple definition, concept, or fact that could be looked up in the textbook (i.e., "what is meant by the polarity of the membrane?").

Category 1b: Questions about a more complex definition, concept, or fact explained fully in the textbook (i.e., "what does it mean when it is says air moves through a bird's lungs?").

Category 2: Ethical, moral, philosophical, or sociopolitical questions (i.e., "carbon monoxide is a very deadly gas binding to hemoglobin much faster than oxygen. If it is so deadly, why are there no carbon monoxide detectors throughout the dorm halls?").

Category 3: Questions for which the answer is a functional or evolutionary explanation. (In this case students begin by asking a question that relates to function and could, in principle, be answered in functional terms—"Why do people have an appendix?"—however, the deeper answer is more often related to evolution than to function (the human appendix is a vestigial organ)).

Category 4: Questions in which the student seeks more information than is available in the textbook (i.e., "what causes the 'rumbling' in your stomach when you are hungry?").

Category 5: Questions resulting from extended thought and synthesis of prior knowledge and information, often preceded by a summary, a paradox, or something puzzling. (i.e., "In chapter 35 it says that caffeine, if taken excessively, can disrupt motor coordination and mental coherence which can cause depression. I known that Coca-Cola has some amount of caffeine in it. Does this mean that excessive consumption of it could lead to depression . . . ?")

Category 6: Questions that contain within them the kernel of a research hypothesis (i.e., "I have heard that some people snore so badly that they stop breathing during their sleep. What correlation is there, if any, between "heavy snorers" and a higher instance of apnea during REM sleep. Can the attention their nervous system is devoting to a dream, interfere the regulation of respiration?").

This taxonomy was developed empirically by examining over 150 written questions from a previous class and grouping similar types of questions together. Armed with "new" taxonomic categories for student questions, we began to examine the distribution of student questions across categories over time—first in the active learning class, and then in the following semester in the traditional class.

Methods

Description of the Two Courses

Biology 100 is a 4-credit, one-semester class designed to introduce students to basic concepts in biology. Over half of the students (ca. 60%) are women, and about 40% are minorities. Most are freshmen (60–75%) or sophomores (15–25%). Half or more are either declared biology and biochemistry majors (30–35%), or are in allied health pre-professional programs (20%). The remaining 45–50%, for whom the course is not a required prerequisite for later courses, include mainly non-science majors, undeclared students, and a few science and engineering majors in other areas.

The *active learning* class employed student-centered, constructivist-based, and interactive instructional approaches widely recommended by nationally prominent educators and scientific organizations (AAAS, 1990; Slavin, 1991; Johnson et al., 1991; Project Kaleidoscope, 1991; Meyers & Jones, 1993; Brooks & Brooks, 1993; NSF, 1996; NRC, 1996, 1997; McNeal & D'Avanzo, 1997). Students met in the large class for three, 50-minute sessions per week, and in ten smaller ($n \sim 25$), weekly discussion sections led by graduate teaching assistants (GTAs). In the discussion section GTAs randomly assigned students in their sections to "teams" (cooperative learning groups of 3–5 students) during the first week of classes, and students were instructed to sit together with their teammates in the large class as well as in their discussion section. In the large class teams were given problems for discussion and subsequent formulation of a written group response. During in-class team exercises—which usually lasted 2–5 minutes—the instructor circulated through the lecture hall listening to team discussions and offering suggestions, comments and answers. After a team exercise, the instructor called on four or five students to share the conclusions of their team. Over the semester, students were encouraged to ask questions about items and issues from assigned reading, from press and broadcasting media reports, from topics covered in lecture, or from personal experience. Questions from students were voiced publicly (using wireless microphones), written in laboratory notebooks, and given to the instructor at the end of a class period, or posed on e-mail directly to the instructor. Student questions were often used to initiate small-group learning exercises and/or to launch whole-class consideration of key biological concepts and processes.

In addition to chapter-based homework exercises (see below) students in the active learning class were asked to write three, short research papers, which we called "take-home assignments" (ThAs), and in each ThA they were asked to pose two original questions based on their research. ThA questions were graded, and counted a third of the overall grade on the assignment. All ThAs were returned to the students with written comments. During the semester, the instructor also discussed with the class what *they* consider to be good questions, brought examples of types of questions that he thought were good questions, and asked students to work with their teammates to rate their own questions (see Sokolove, 1998 for an extended description of the active learning class).

The *traditional class* was taught using a lecture format with little time allocated for open discussion. Students rarely asked questions during lectures and there were only a few students who participated more than once or twice over the course of the semester. In this class there were no wireless microphones and it was difficult for the students to hear each other's questions. Students rarely asked the instructor content-related questions in notes or on e-mail. Also, unlike students in the active learning class, students in the traditional class were not asked to write research papers (i.e., ThAs). The main opportunity for students to ask "real" (i.e., their own) questions was in chapter-based homework exercises, which were neither graded nor returned.

Description of Homework Exercises in the Active Learning Class

Four times during the semester students were asked to formulate a question after reading a chapter from their textbook (Starr & Taggart, 1998). In each case one, or two closely related chapters were assigned for reading, and students were directed to bring a typewritten question to hand in at the start of the following class period. No "extra credit" points were given. Such assignments were credited toward the final grade (along with numerous other hand-ins) as "non-verbal participation."

HW0. The first chapter assigned dealt with cell energetics (Chapter 8). Each student brought his/her question to the class, sat with teammates, and followed the instructions to select the "best" question that had been formulated by a member of the team. Because the instructor collected only the "best" questions chosen by each team and not *all* of the questions that students wrote, we have not included results of this assignment in the current report. Reference to it is made here in order to show what students were asked to do.

HW1. The second chapter assigned dealt with meiosis (Chapter 10). The instructor collected all of the questions brought in by the students. There was no team selection of "best" questions.

HW2. The third assignment was based on two chapters (Chapters 13 and 14), the first of which dealt with DNA structure and function and the second with the relation between DNA and proteins. Students brought their questions to class and sat with their teammates to choose the best question that had been formulated by a member of their team. Best questions were placed on the top of each team's stapled stack. The instructor collected all of the questions and kept each team's questions in the same order that the students had arranged them.

HW3. The fourth assigned chapter dealt with digestion (Chapter 42). This exercise was given to the class toward the end of the semester. Once again students brought their questions to class and sat with their teammates to choose the best question that had been formulated in their team. As in HW2 above, the best questions were placed on the top of each team's stapled stack. The instructor again collected all of the questions and kept them in the same order that the students had arranged them.

After the third chapter-based homework exercise (HW2) the instructor presented the new taxonomy in class and provided an example question for each of the categories. The instructor noted he considered questions in categories 5 and 6 to be the best type of questions for a science course.

Description of Homework Exercises in the Traditional Course

Five times during the semester students were asked to formulate a question after reading a chapter in the same textbook (Starr & Taggart, 1998). More exercises were included in the traditional than in the active learning class in part to compensate for the absence of other formal, question-asking opportunities such as ThAs. The instructor also offered five "extra credit" points for each question as an additional incentive to complete the assignment. In the traditional class (in contrast to the active learning class), only exams counted toward the final course grade and each completed homework exercise was roughly equivalent in value to 1.5 midterm exam questions. It was important to encourage these students, because unlike the students in the active learning class they were not used to doing such assignments.

HW1. The first homework exercise for the traditional course dealt with a chapter on DNA structure and function (Chapter 13). This exercise was given to the students near the beginning of the semester. In the active learning class (see above) Chapter 13 was also one of two chapters assigned in HW2.

HW2. The second homework exercise was based on the reading of two chapters (Chapters 17 and 18) dealing with evolution.

HW3. The third homework exercise was based on the reading of two chapters (Chapters 34 and 35) dealing with the nervous system. This exercise was given to the students toward mid-semester.

HW4. The fourth chapter dealt with digestion (Chapter 42). The exercise was given to the students toward the end of the semester, at about the same time that this same chapter had been assigned to the "active learning" class.

HW5. The fifth chapter dealt with respiration (Chapter 41). This chapter was selected as the "post-taxonomy" chapter, because the topics covered in this chapter were similarly physiological (vertebrate respiration), and our prior experience was that students are equally interested in both topics—probably because in both cases they can see a direct connection to themselves and their own bodies.

Students handed in their questions at the beginning of the class period following the one in which the assignment had been given, and there was little discussion about the assignment questions. However, after the fourth homework exercise (HW4) the instructor presented the new taxonomy in class, and provided clear example questions for each category. All information was provided in handouts as well as verbally. As he had in the active learning class, the instructor emphasized that he preferred questions that fell into categories 5 or 6.

Statistical Analysis

To compare the distribution of sets of student questions produced in response to different assignments, we employed two methods. The first yielded a rough estimate of differences between sets of questions. The second provided a non-parametric test of statistical significance.

To get an impression of the general similarities and differences between question sets we arbitrarily assumed there was a single unit of "question type" separating each taxonomic category and we converted the eight categories into eight cardinal numbers (9–16) which were multiplied by the number of the question in each category. The average of the sum of these products was taken as the "mean" of the distribution. By comparing the mean value we could easily and rapidly assess whether the distribution means were widely separated, and if so, in which direction. However, because the set of questions distributes across categories in a non-normal fashion it was necessary to compare question sets non-parametrically (i.e., *t*-tests are inappropriate). Comparison of question sets across successive assignments over the semester was accomplished using the Wilcoxon Signed Ranks Test provided in the SPSS statistical analysis package.

In order to compare improvement in question-asking by academic major, gender, race/ethnicity and SAT score we looked only at students who turned in both the first and the last homework assignments. First, we verified that the distribution of those who completed both assignments was similar to the distribution of subgroups in the whole class. We then used the Phi coefficient to evaluate the difference between the frequency of students who improved and those who did not improve within the subgroups. The Phi coefficient compares the observed values for each subgroup with the expected frequencies derived from the proportion of subgroup members within the population of students who turned in both exercises.

Results

Below we describe separately our findings regarding the traditional class and the active learning class beginning with a summary of our findings in the active learning class, including data about the relationship between improvement of students questions within different student sub-populations (academic major, gender, race/ethnicity, and combined SAT scores).

The Active Learning Class

Table 2 summarizes findings in the active learning class (Marbach-Ad & Sokolove, in press).

The distribution for HW1 questions was similar to the distribution for HW2 questions except in categories 4 and 5. In HW1, more students (23%) formulated questions that fell in category 5 (questions resulting from extended thought and synthesis) than they did in HW2 (13%). In contrast, in HW2 many more students (32%) formulated questions that fell in category 4 (questions requesting specific information not readily found in the textbook) than they did in HW1 (20%). Because chapters 13/14 (HW2) dealt with topics such as mutation and genetic diseases, it may be that students were more interested in obtaining additional information (i.e., ‘‘how can DNA become damaged?’’) than in posing thoughtful questions.

The distribution of questions in HW3 (after the new taxonomy was presented and discussed in class) was distinctly different from the distribution for student questions in either HW1 or

Table 2
Percentage of students' questions in each category in homework exercises in the active learning class

Category	HW1 $N = 182$	HW2 $N = 188$	HW3 $N = 173$
(0) Based on misunderstanding	15%	13%	8%
(1a) Simple definition question	12	11	5
(1b) Complex definition question	12	11	9
(2) ‘‘Motives’’ or ‘‘intentions’’	1	3	2
(3) Evolutionary questions	10	10	2
(4) Seeking more information	20	32	30
(5) Thoughtful questions	23	13	30
(6) Research questions	7	7	14

HW2. In HW3 in each of the "lower-level" Categories (Categories 0–3) the percentage of questions was notably lower than the percentage of these questions asked in HW1 or in HW2 (Table 2 and Fig. 2; categories 0–3 are not all low-level categories, but we have grouped them together here in order to focus on categories 4–6). The percentage of questions that fell into category 4 did not change much from HW2 (32%) to HW3 (30%). However, in HW3 there were twice as many questions that fell into categories 5 (30%) and 6 (14%), than in HW2 (category 5, 13%; category 6, 7%).

Non-parametric analysis showed no significant difference between HW1 and HW2. In contrast, the difference between the set of questions in HW1 (or HW2) and HW3 was highly significant ($p < .001$).

We have also examined the improvement in the nature of questions from the first exercise (HW1) to the last exercise (HW3) across different sub-populations in the class. In each case (academic major, gender, race/ethnicity, and combined SAT scores) we have compared the percent of students from a sub-population who showed improvement and the percent of the sub-population in the class as a whole (Table 3). Thus, for example, 62% of all students whose questions improved were female and the fraction of female students in the entire class was 60%. Table 3 shows that in the active learning class students with higher SAT scores did improve somewhat more than those with lower SAT scores (although the improvement was not statistically significant), but other factors such as academic major, gender, race/ethnicity do not appear to play a role in the improvement in students' questions over the course of the semester.

Table 3

Sub-populations of students in traditional and active learning classes who improved their questions compared to their distribution in the whole class

	Active Learning Class		Traditional Class	
	Percent of Students who Improved[a]	Percent of Students in the Class	Percent of Students who Improved[b]	Percent of Students in the Class
Primary major				
Biology or Biochemistry	35%	32%	41%	33%
Other and undecided	65	68	59	67
Gender				
Female	62	60	62	62
Male	38	40	38	38
Ethnicity[c]				
Caucasian	59	60	49	55
Asian	24	25	15	22
African American	13	14	31	22
Other	4	1	5	1
Combined SAT score				
Bellow the Median	41	49	44	43
Above the Median	51	42	41	44
Unknown	8	9	15	13

[a] The number of students who improved in the active learning class = 71 out of 131 who did both HW1 and HW3.

[b] The number of students who improved in the traditional class = 39 out of 87 who did both HW1 and HW5.

[c] The only significant difference was in the traditional class where the populations of those who improved differed from their corresponding proportions in the whole class ($\phi = .54$; $p < .001$).

Table 4
Percentage of students' questions in homework exercises in the traditional class

Category	HW1 $N = 120$	HW2 $N = 101$	HW3 $N = 117$	HW4 $N = 108$	HW5 $N = 123$
(0) Based on misunderstanding	22%	27%	23%	25%	27%
(1a) Simple definition	15	9	21	17	8
(1b) Complex definition	5	14	3	4	9
(2) Motives or intentions	1	1	0	1	3
(3) Evolutionary questions	14	2	9	11	3
(4) Seeking more information	30	32	35	33	33
(5) Thoughtful questions	10	13	9	7	9
(6) Research questions	3	2	0	2	8

The Traditional Class

Table 4 shows the results from the traditional class. All five homework exercises exhibited a similar pattern with only minor variations.

In all five exercises about a quarter of the questions (22–27%) fell into category 0 (questions do not make logical or grammatical sense, or are based on a basic misunderstanding or misconception). In HW5, after the instructor presented the taxonomy to the class and emphasized that he preferred research questions, we identified some questions that seemed to be "research questions", but actually belonged in the 0 category because they were based on a misconception (i.e., "is it possible to increase the levels of myoglobin in the skeletal muscle, perhaps by experimentally introducing it into the body, or somehow manipulating the body to produce more?"). Particularly in HW5 it seemed that some students in the traditional class tried hard to pose "good" questions, but it was beyond their ability and/or their knowledge base to do so.

When we combined categories 1a and 1b, we found that roughly 20% of the questions fell into the combined category in all five exercises with a slight decrease in HW5 (20, 23, 24, 21, 17%).

Generally there were only a few questions from category 2 (questions that dealt with ethical, moral, philosophical, or sociopolitical issues). In the active learning class this type of question was sometimes the basis of a whole-class discussion, but in the context of the assigned homework it was not deemed as desirable or as oriented to science process as those in categories 4–6.

Questions that fell into category 3 (questions seeking information about why the world is the way it is) seemed to be topic-dependent. Their frequency varied widely (range 2–14%), but did not appear to be related to time in the semester (14, 2, 9, 11, 3%). In all five exercises about a third of the questions (30–35%) fell into category 4, regardless of chapter topic or time in the semester.

In all five exercises, about 10% of the questions (7–13%) fell into category 5. In category 6 we observed what appeared to be a small but definite improvement in the final exercise (HW5) after the new taxonomy had been presented to the class (3, 2, 0, 2, 8%).

Non-parametric analysis showed no significant differences among the five exercises. Although there were no significant differences between the overall improvement in students' questions over the semester, we also (as in the active learning class) investigated the type of students who improved their questions from the first to the last exercise. The results indicate that factors such as academic major, gender, and combined SAT scores are not important factors in the improvement of questions in the traditional class (Table 3). However, when we examined question-improvement in students of different racial or ethnic origins we did find a significant difference ($p < .001$). Although the numbers are small, the results indicate that in the traditional class more improvement occurred in African Americans than in other ethnic groups.

Summarizing the Findings for Both Classes

In order to view the overall responses in the two classes, in each class we combined categories 0–3, which included generally less sophisticated or desirable (from the science instructor's perspective) questions, and categories 4, 5, and 6, which were more desirable, sophisticated, thoughtful, and/or insightful.

Category 4 questions represent the type of curiosity-based scientific inquiry that serves as a driver of the scientific process. They are thus important and desirable in the context of a college-level science class, and it is appropriate to group them together with higher level category 5 and 6 questions. Questions in categories 5 and 6 are science-related questions that go beyond simple curiosity about the world. They are questions in which students must employ higher level thinking skills such as synthesis and integration of information acquired at different times in the semester. The distinguishing feature of a category 6 question is that it contains within it the kernel of an hypothesis. (Although category 1 questions are also curiosity-based, unlike category 4, they do not go beyond the framework of the textbook.)

In the active learning class [Fig. 1] in HW1 and HW2 about half of the questions fell in categories 0–3 and the other half fell into categories 4–6. However, in HW3 only 26% of the questions fell in categories 0–3 while 76% fell into the higher numbered categories. Similar improvement was seen in research papers (take-home assignments or ThAs) where students were instructed to pose questions (see Methods, and Ad-Marbach & Sokolove, 1999). In contrast, in the traditional class [Fig. 2] roughly half of the questions in all five assignments fell into categories 0–3 and the other half fell into categories 4–6.

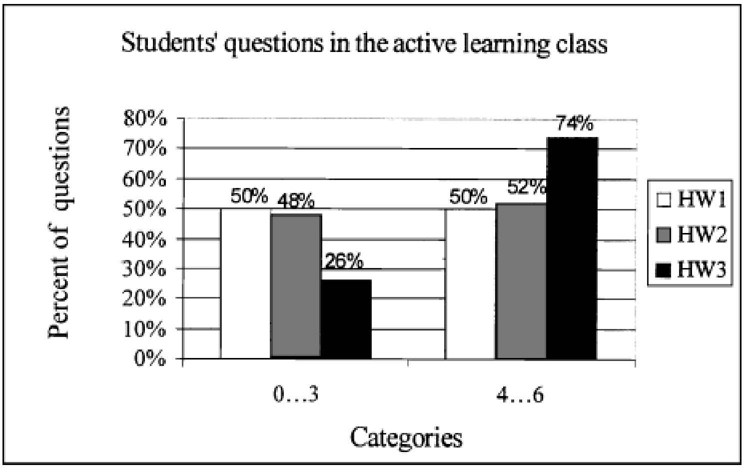

Figure 1. Distribution of student questions in an active learning class written after reading one or more textbook chapters.

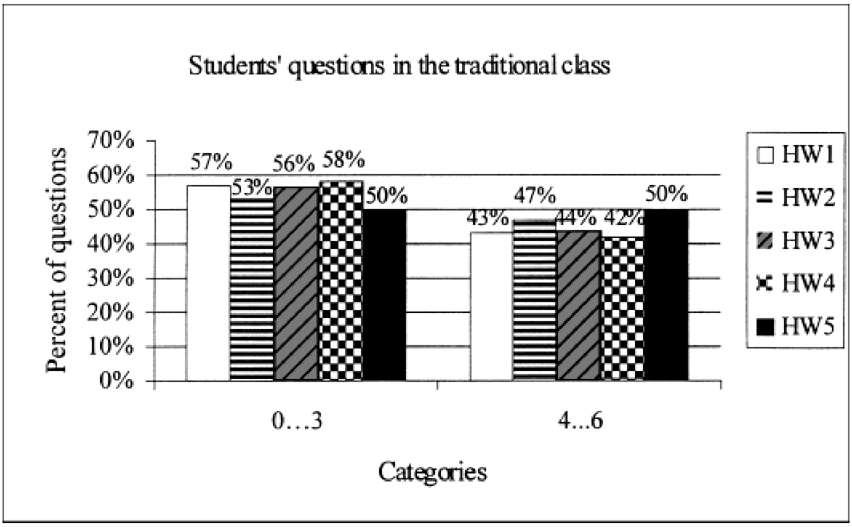

Figure 2. Distribution of student questions in the traditional class written after reading one or more textbook chapters.

Discussion

This study is part of a longitudinal comprehensive research project that will examine whether exposure to student-centered, active-learning classes in introductory biology makes a difference across multiple years with respect to students' achievement, attrition, and attitudes and beliefs toward and about the nature of science. Introductory biology classes tend to be large lecture courses, and this generally tends to reinforce students' roles as passive learners who absorb concepts and facts only long enough to get through the next test. Thus, it is a special challenge for such courses to promote biological understanding and self-confidence in students (Ebert-May, Brewer, & Allred, 1997).

An obvious feature of our active learning class is that because students speak in class using wireless microphones, both the instructor and the other students in the class can hear the person speaking. Thus, students can ask questions during the lesson, the instructor and/or other students can easily and naturally respond to them, and question-asking becomes a normal component of student behavior in the lecture hall. A second feature is that in an active learning environment with cooperative learning groups the instructor can ask students to formulate questions, to work with their teammates and decide in class which of their questions are the best ones, and to share their thoughts immediately with the whole class.

The active learning class was offered in Fall 1998, one semester before the traditional class. When we saw the improvement in questions in the active learning class, we postulated that the observed change could have been due to one or more of the following:

1. The time in the semester without any connection to teaching style. Toward the end of the semester, students would be expected to gain more content knowledge from this course (and perhaps from other courses in which they were enrolled), and therefore they would begin to ask better questions.
2. The fact that before HW3 the instructor had presented the new taxonomy in class and provided example questions for each category. This intervention alone might have

resulted in improved questions, because it helped the students better understand the type of questions the instructor desired.

3. The fact that students had experienced an active learning environment and had been exposed to many different assignments in which they were asked to formulate questions and to identify "good" questions.

The results in the traditional course contradict the first assumption that the improvement in students' questions in the active learning class was due to the later time in the semester. In both classes students had been asked to pose a question after reading the same chapter (Chapter 42) at about the same time in the semester (toward the end). In the traditional class less than half of the students posed higher level questions after reading chapter 42 [HW4, Fig. 2], while in the active learning class about three-quarters of the class posed higher level questions based on Chapter 42 [HW3, Fig. 1].

The results from the traditional course also contradict the second assumption that the presentation of the new taxonomy in class was the primary cause for the improvement of questions in the active learning class. In the traditional class we presented the new taxonomy to the students after HW4 and before HW5, and we saw only a minor improvement in the questions that students asked in HW5.

In summary, there was a significant improvement in students' questions toward the end of the semester in the active learning class following a specific intervention, while in the traditional class there was little, if any, improvement following the same intervention. By many measures (gender, minorities, etc.; see Methods) both sections comprised similar populations, and it therefore seems reasonable that they should have similar potential in their ability to formulate questions. We believe that the difference in outcomes is most likely due mainly to the differences in pedagogy (including but not limited to the difference in emphasis on question-asking) between the classes (Table 1).

It is difficult to say which of the many differing pedagogical methods detailed in Table 1 were the most influential in eliciting improvement of students' questions in the active learning class. (For example, in the active learning class choosing the "best" question from each group was an exercise that involved participation of all group members in the decision about which question *was* the "best" question.) We suspect that it was, in fact, a mixture of all of them together. Although in-class presentation of the taxonomy influenced question-asking in both classes, substantive improvement was seen in the class that encouraged question-asking and evaluation of questions in multiple instances across the semester. Thus, a large improvement was seen in the active learning class (categories 5 and 6, Table 2), whereas in-class presentation of the taxonomy in the traditional class, produced only a small improvement in students' questions (category 6, Table 4).

Conclusions and Implications

The value of an active learning class that includes cooperative learning groups has been highlighted by theoreticians, researchers, and practitioners mainly in K-12, but also in college (Radebaugh & Kazemek, 1989). Most researchers have focused on academic achievement (Slavin, 1983), although a few have focused on students' questions. In our introductory biology class we have focused on students' questions for two reasons: First, we believe that one way in which teachers can promote the notion of independence in learning includes the provision of opportunities for students to become questioners (Woodward, 1992), and second, we believe simply (along with virtually all of our science colleagues) that good science begins with good questions.

We are aware that one of the criticisms of "active learning" methods is that the instructor cannot usually manage to teach the same amount of subject matter or cover all the content that he or she might cover in a traditional course. However, we feel that content-coverage is not the *only* important objective in teaching. In 1944, John Dewey, noted that "all which the school can or need do . . . is to develop [students'] ability to think (p. 152)." "Real" questions that seek more information require students to think. They must think about what it is that they already know and understand and identify what it is that they do not know or understand.

No one can argue that it is easy to teach large classes in college in an active learning mode. However, conclusions from our experiment suggest that even in *traditional* classes students can improve their study skills when presented with multiple opportunities to ask questions. One of the major complaints of instructors is that "students do not read the book before they come to class." Focusing on question-posing forces students to read the text first in order to write their questions. In Biol 100 we deliberately asked students to type their questions so that they would be more likely to complete the assignment at home and not during the first minute or two of class.

Regardless of whether one is teaching in a traditional or an active learning class, students' questions provide an opportunity to understand what students have in mind and help to uncover their misconceptions and/or preconceptions. However, it is especially important to obtain and read such questions in the traditional course where students rarely have an opportunity to express their thoughts, and where most students rarely ask even a single question over the course of a semester. Yager (1991) asks, "How can science teachers move towards constructivist approaches? (p. 56)", and the first procedure that he describes is: "seeking out and using student questions and ideas to guide lessons and whole instructional units (p. 56)."

At least four categories in our taxonomy can be particularly helpful in probing students' understanding. Category 0 questions help to identify the type of preconceptions or misconceptions students have after reading a textbook chapter at home. Most of the category 1a questions suggest that the students who wrote them did not read the chapter at all, or did not read it carefully, while most of the category 1b questions suggest that the students who wrote them did read the chapter carefully, but did not understand what they read. Category 4 questions (those that seek more information than can be found in the textbook) help identify what students are most curious about or interested in.

Based on our experience and the data from this study, we suggest first, that teachers present the student–question taxonomy to students at the *beginning* of the semester to let them know what is expected. Second, even if an instructor feels uncomfortable about teaching in active learning style (or, as may be more often the case, feels unable to cover enough content material in an active learning classroom), he/she should try as much as possible to encourage and respond to students' questions. When one does so, students instinctively begin to recognize and appreciate the importance of questions in the science classroom: "Even with such a large class, I think Dr. Sokolove does an excellent job getting to everyone's questions and comments" (from an anonymous mid-point student feedback form given to students in the *traditional* course).

Research supported in part by a grant from the National Science Foundation, REC-9815007. Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the National Science Foundation.

The authors thank Ellen Beck and Dr. Marina Adler, Department of Sociology, UMBC, for help in processing the data and providing statistical analysis.

References

Ad-Marbach, G., & Sokolove, P.G. (1999). A new taxonomy for student questions. Paper presented at the Lilly Conference on Teaching and Learning—Atlantic, Towson, MD, April 16–18.

American Association for the Advancement of Science (AAAS) (1990). Science for all Americans. New York: Oxford University Press.

Bean, T.W. (1985). Classroom questioning strategies: Directions for applied research. In A.C. Graesser & J.B. Black (Eds.), The psychology of questioning (pp. 335–358). Hillsdale, NJ: Erlbaum.

Bloom, B.S. (Ed.) (1984). Taxonomy of educational objectives: Handbook 1: Cognitive domain. New York: Longman Inc.

Brooks, J.G., & Brooks, M.G. (1993). In search of understanding: The case for the constructivist classroom. Alexandria, VA: Association for Supervision and Curriculum Development.

Caprio, M.W. (1993). Cooperative learning—the jewel among motivational-teaching techniques. Journal of College Science Teaching, 22, 279–281.

Commeyras, M. (1995). What can we learn from students' questions? Theory into Practice, 34(2), 101–106.

Cooper, M.M. (1995). Cooperative learning: An approach for large enrollment courses. Journal of Chemical Education, 72 (2), 162–164.

Cooper, J., Prescott, S., Cook, L., Smith, L., & Mueck, R. (1990). Cooperative learning and college instruction. California: California State University.

Dewey, J. (1944). Democracy and education. New York: The Free Press.

Dillon, J.T. (1988). The remedial status of student questioning. Journal of Curriculum Studies, 20(3), 197–210.

Dori, Y.J. & Herscovitz, O. (1999). Question-posing capability as an alternative evaluation method: Analysis of an environmental case study. Journal of Research in Science Teaching, 36, 411–430.

Ebert-May, D., Brewer, C., & Allred, S. (1997). Innovation in large lectures—teaching for active learning. BioScience, 47 (9), 601–607.

Good, T.L., Slavings, R.L., Harel, K.H., & Emerson, H. (1987). Student passivity: A study of question asking in K-12 classrooms. Sociology of Education, 60, 181–199.

Heady, J.E. (1993). Teaching embryology without lectures and without traditional laboratories—An adventure in innovation. Journal of College Science Teaching, 23 (2), 87–91.

Houston, V.M. (1938). Improving the quality of classroom questions and questioning. Educational Administration and Supervision, 24, 17–28.

Johnson, D.W., & Johnson, R.T. (1994). Learning together and alone: Cooperative, competitive, and individualistic learning. Massachusetts: Allyn and Bacon.

Johnson, D.W., Johnson, R.T., & Smith, K.A. (1991). Active learning: Cooperation in the college classroom. Edina, MN: Interaction Book Co.

Kagan, S., & Roberson, L. (1992). Co-op facilitators' handbook. San Clemente, CA: Kagan Cooperative Learning.

Marbach-Ad, G., & Sokolove, P.G. (in press). Good science begins with good questions. Journal of College Science Teaching.

McNeal, A.P., & D'Avanzo, C. (Eds.) (1997). Student-active science: Models of innovation in college science teaching. New York: Harcourt Brace.

Meyers, C., & Jones, T.B. (1993). Promoting active learning: Strategies for the college classroom. San Francisco: Jossey-Bass.

National Research Council (NRC). (1996). National science education standards. Washington, DC: National Academy Press.

National Research Council (NRC). (1997). Science teaching reconsidered, A handbook. Washington, DC: National Academy Press.

National Science Foundation (NSF). (1996). Shaping the future: New expectations for undergraduate education in science, mathematics, engineering, and technology. Washington, DC: National Science Foundation.

Olson, G.M., Duffy, S.A., & Mack, R.L. (1985). Question-asking as a component of text comprehension. In A.C. Graesser & J.B. Black (Eds.), The psychology of questioning (pp. 219–226). Hillsdale, NJ: Erlbaum.

Orr, H.A. (1999). An evolutionary dead end? Science, 285, 343–344.

Pearson, J.C., & West, R.L. (1991). An initial investigation of the effects of gender on student questions in the classroom: Developing a descriptive base. Communication Education, 40, 22–32.

Pechenik, J.A. (1993). A short guide to writing about biology. 2nd ed. New York: Harper Collins College Publishers.

Project Kaleidoscope (1991). What works: Building natural science communities. Vol. 1. Washington, DC: Stamats Communications, Inc.

Radebaugh, M.R., & Kazemek, F.E. (1989). Cooperative learning in college reading and study skills classes. Journal of Reading, 32 (5), 414–418.

Resnick, L.B. (1987). Learning in school and out. Educational Researcher, 16, 13–20.

Slavin, R.E. (1983). When does cooperative learning increase student achievement? Psychological Bulletin, 94, 429–445.

Slavin, R.E. (1991). Synthesis of research on cooperative learning. Educational Leadership, 48, 71–82.

Sokolove, P.G. (1998). The challenge of teaching Biology 100: Can I really promote active learning in a large lecture? In M.B. Gardner & D.L. Ayres (Eds.), Journeys of transformation (pp. 121–128). College Park, MD: Maryland Collaborative for Teacher Preparation.

Starr, C., & Taggart, R. (1998). Biology: The unity and diversity of life. 8th ed. Belmont, CA: Wadsworth Publishing Company.

Watts, M., Gould, G., & Alsop, S. (1997). Questions of understanding: Categorizing pupils' questions in science. School Science Review, 79(286), 57–63.

West, R., & Pearson, J.C. (1994). Antecedent and consequent conditions of student questioning: An analysis of classroom discourse across the university. Communication Education, 43, 299–311.

Wood, D., & Wood, H. (1988). Questioning versus student initiative. In J.T. Dillon (Ed.), Questioning and discussion (pp. 280–305). Norwood, NJ: Albex.

Woodward, C. (1992). Raising and answering questions in primary science: Some considerations. Evaluation and Research in Education, 6, 145–153.

Yager, R.E. (1991). The constructivist learning model. The Science Teacher, 58 (6), 52–57.

Zinsser, W. (1976). On writing well. New York: Harper and Row, Publishers.