

Article

Peer Learning and Support of Technology in an Undergraduate Biology Course to Enhance Deep Learning

Masha Tsaushu,* Tali Tal,* Ornit Sagy,* Yael Kali,[†] Shimon Gepstein,[‡] and Dan Zilberstein[‡]

*Department of Education in Technology and Science and [†]Faculty of Biology, Technion–Israel Institute of Technology, Haifa 32000, Israel; [‡]Technologies in Education Graduate Program, Department of Learning, Instruction and Teacher Education, Faculty of Education, University of Haifa, Mount Carmel, Haifa 31905, Israel

Submitted April 12, 2012; Revised August 20, 2012; Accepted September 1, 2012
Monitoring Editor: Clarissa Ann Dirks

This study offers an innovative and sustainable instructional model for an introductory undergraduate course. The model was gradually implemented during 3 yr in a research university in a large-lecture biology course that enrolled biology majors and nonmajors. It gives priority to sources not used enough to enhance active learning in higher education: technology and the students themselves. Most of the lectures were replaced with continuous individual learning and 1-mo group learning of one topic, both supported by an interactive online tutorial. Assessment included open-ended complex questions requiring higher-order thinking skills that were added to the traditional multiple-choice (MC) exam. Analysis of students' outcomes indicates no significant difference among the three intervention versions in the MC questions of the exam, while students who took part in active-learning groups at the advanced version of the model had significantly higher scores in the more demanding open-ended questions compared with their counterparts. We believe that social-constructivist learning of one topic during 1 mo has significantly contributed to student deep learning across topics. It developed a biological discourse, which is more typical to advanced stages of learning biology, and changed the image of instructors from "knowledge transmitters" to "role model scientists."

INTRODUCTION

Criticism of teaching in higher education institutions has been growing in recent years. Among international scientists, there is increasing agreement about the need to change the culture of science education in research universities to promote more meaningful learning. This change will only happen when we find the balance between technical interests in science and human interests in science learners (Gilmer, 2010). Anderson and his colleagues, biomedical research scientists represent-

ing a diversity of institutions, argued that to maintain the vitality of research universities requires a culture in which teaching and research support two equally important enterprises: generation of new knowledge and education of students (Anderson *et al.*, 2011). Both scientists and science educators have called for a program that refrains from merely providing broad content. They have emphasized the need to develop students' analytical skills, while promoting understanding of scientific research processes and inspiring curiosity (Gilmer, 2010).

In 2009, a conference hosted by the American Association for the Advancement of Science, with support from the National Science Foundation, brought together faculty, administrators, students, and other educational stakeholders to discuss biology teaching at the undergraduate level. The recommendations from the meeting can be summarized as "the biology we teach should be the biology we do," meaning that in addition to learning the content, students should gain a better understanding of the nature of science and that assessments should help instructors to figure out how deeply students understand (or misunderstand) the basics of the

DOI: 10.1187/cbe.12-04-0042

Address correspondence to: Masha Tsaushu (tmasha@gmail.com).

© 2012 M. Tsaushu *et al.* CBE—Life Sciences Education © 2012 The American Society for Cell Biology. This article is distributed by The American Society for Cell Biology under license from the author(s). It is available to the public under an Attribution–Noncommercial–Share Alike 3.0 Unported Creative Commons License (<http://creativecommons.org/licenses/by-nc-sa/3.0>).

"ASCB[®]" and "The American Society for Cell Biology[®]" are registered trademarks of The American Society for Cell Biology.

discipline, rather than test for recall of facts or repetition of memorized procedures. In addition, Web- and print-based tools should be available to help students access and interact with the information, which will enable them to develop tools to acquire understanding and to become part of a scientific community (Woodin *et al.*, 2009).

Large enrollments have always been typical of introductory undergraduate courses, but the growth of universities worldwide and the global economic crises have further worsened the faculty–student ratios (Haak *et al.*, 2011). Consequently, most introductory courses rely on lectures that attempt to “deliver the content,” a technique that has proven to be ineffective in fostering conceptual understanding of scientific reasoning (Handelsman *et al.*, 2004). As the primary means of informing students, it can hardly inspire curiosity or motivate learning (Gilmer, 2010). The knowledge introductory courses attempt to teach is constantly growing, especially in biology, and the need to change the teaching philosophy is therefore even more acute. If we want to highlight and discuss main principles and complex ideas, expose students to biological research, and encourage “biological thinking,” we cannot do it with a professor who struggles to teach the entire textbook to 300 individuals while standing on a distant podium in a large lecture hall.

The scientific community has not ignored the above-mentioned challenges, and there are ongoing attempts to suggest other models. Universities and institutions, such as the U.S. National Institutes of Health (NIH), execute programs that promote effective pedagogical approaches to undergraduate education in biology (Woodin *et al.*, 2010). Programs such as the 1-wk NIH Summer Institute, which has taken place annually since 2004, have shown a positive multidimensional impact on the participants’ teaching methods (Pfund *et al.*, 2009). Many attempts to supplement or replace science, technology, engineering, and mathematics (STEM) lectures with active learning have been made in the past decade, increasing student conceptual understanding and improving their attitudes toward these courses (Henderson *et al.*, 2011).

The teaching effort and research described in this paper took place in a major research university in Israel. It was a result of a discussion between two biology professors who were dissatisfied with their teaching and a group of four science educators. Together, we developed and implemented this innovative instructional model. In line with Handelsman *et al.* (2004) and Woodin *et al.* (2009), we aimed to make teaching more scientific and student learning more active and meaningful in the large-enrollment introductory course Biology 1. Our model gives priority to resources not used enough to enhance active learning in higher education: technology and the students themselves. Given the constraints of the class size and a syllabus that we could not change, we substantially reduced the number of lectures and replaced them with educational technology that supported individual learning and short-term, small-group learning.

This study is framed with the view of learning as a social-constructivist activity, as well as a cognitive process, that can take place face-to-face or through online interactions (Linn and Hsi, 2000). The social-constructivist approach that has developed from Vygotsky’s theory and through scholars who followed him emphasizes critical dialogue with the teacher or among peers to promote meaningful learn-

ing (Driver *et al.*, 1994; Ash, 2004). Vygotsky also argued for an essential distinction between scientific and common conceptions. The notion, for example, that mushrooms and humans share the same basic mechanisms is certainly not a common-sense assumption (Klymkowsky *et al.*, 2003). This difficulty in scientific understanding of the natural world is one of the main factors that inhibit the development of a deep approach to science learning, especially among young people.

The distinction between deep and surface learning shows that students who use a surface approach give “black box” explanations that do not refer to mechanisms and tend to ask about more basic, factual, or procedural, information (Chin and Brown, 2000), while students who use a deep-learning approach give more elaborate explanations that describe mechanisms and cause–effect relationships; ask questions that focus on explanations, causes or predictions; and engage in “online theorizing” (Marton and Saljo, 1976). The idea of a deep versus a surface approach was recently discussed by Gilmer (2010), who studied her own shift in practice when teaching college biochemistry using technology and small-group learning.

Marton and Saljo showed that students adapt their way of learning to their conception of what is required of them. This phenomenon was more recently documented by Scouller and Prosser (1994). Deep-learning strategies were employed by students preparing their assignment essays; these strategies were perceived by the students as application of higher levels of cognitive processing. In contrast, surface-learning approaches employed by the students in a multiple-choice (MC) exam context were apprehended by the students as merely knowledge-based (Scouller, 1998).

In light of the above, we addressed the calls to change the style of lecturing in large introductory science classes in higher education, assuming that a pedagogical change informed by the social-constructivist approach would affect the depth of student learning.

This study aims to investigate the impact on the students’ learning of the instructional model we developed and implemented.

On the basis of the instructors’ impression with respect to students’ questioning in class, and in line with the literature that views student questioning as an opportunity to express higher-order thinking (Ennis, 1987; Dori and Herscovitz, 1999; Marbach-Ad and Sokolove, 2000; Hofstein *et al.*, 2005), we were also interested in how the change in the instructors’ focus affected the type and depth of students’ questions in class.

The research question we followed was: How did the instructional change affect learning as revealed by:

1. Questions students asked during lectures, and
2. Students’ achievements in tasks requiring various thinking skills?

Despite calls to change common assessments in higher education, achievement tests are widespread in undergraduate studies, particularly in large-enrollment courses. Nevertheless, it is important to note that we measured student performances in simple and complex items and in closed and open-ended ones. The different types of questions reflected

Table 1. Design of the instructional change

	Preintervention		Intervention	
	Traditional	Traditional plus tutorial	Adapted teaching	Active learning
Tutorial	–	+	+	+
Lectures	+	+	+ Adapted: focus on chosen wide-scope issues	Reduced to 30%
Small-group learning	–	–	–	+
Learning pattern	Passive only	Passive and individual–interactive (with tutorial)	Passive and individual–interactive (with tutorial)	Active individual and group; interactive (with tutorial)
Student assessment	Final MC exam	Final MC exam and O-HOT questions	O-HOT questions	Final MC exam and O-HOT questions; team’s presentation

different thinking skills, as we elaborate in the following sections.

METHODS

The Instructional Change

Institutional policy and constraints ruled out changing the syllabus or the number of students enrolled in the course. Nor could we affect student attendance in class, as it is not obligatory.

Moving away from the traditional lecture format of the course, the instructional change was designed to be implemented in three consecutive phases to allow gradual change and step-by-step, follow-up student learning. In fact, as will be explained later, due to the university’s constraints, some of the versions of the course were carried out simultaneously. Hereafter, we use the term “versions” to describe the different teaching approaches in what was planned to be sequential phases. The design of the instructional change is presented in Table 1.

Traditional-plus-Tutorial Version. Regular lectures existed, as in the traditional format, in which the instructor taught the entire syllabus. In addition, a tutorial was developed and placed on the course website to support independent learning. The tutorial consisted of the videotaped lectures synchronized with PowerPoint presentations and allowed the students to move back and forth. Interactive visualizations, self-feedback questions, a glossary, and discussion forums were incorporated in the online tutorial as well. The purpose of this version was to improve the tutorial, which was expected to support learning in the further versions in response to students’ feedback, and to examine the nature of its usage.

Adapted-Teaching Version. In this version we began changing the teaching approach, asking the instructor to reduce the time dedicated to teaching informative topics and to focus more on complex topics and ideas that have special significance and implications. The instructor, together with M.T., who is a science educator, chose the complex subjects to be discussed in class. Then, while teaching, the instructor emphasized the integration between topics and ideas and highlighted connections between theory, contemporary research, and innovations in biology and biomedicine. Because some of the informative content was no longer presented in class, the students were directed to use the online tutorial, which

included the lectures as videotaped before the intervention, and the textbook to prepare for class.

Active-Learning Version. The number of lectures was reduced to 30% of the number before intervention, comprising mainly an opening and a wrap-up for the course. The students were directed to use the time saved for independent learning, using the online tutorial.

The main innovative component of this version was a group study of one topic. Each group focused on one of the main topics from the course syllabus. The groups were then divided into teams that were guided by teaching assistants (TAs). Each team studied one subtopic in depth for a month and concluded by preparing and presenting it to the entire group. Table 2 presents an example of group topics and subtopics assigned to different teams.

The active-learning version was designed for 300 students per semester. We planned to divide them into 10 groups, each of about 30 students focusing on one of the topics from the course syllabus. Each group would then divide into five teams of six students mentored by a TA for 1 mo. Overall, five TAs were needed for the entire cohort. Figure 1 presents a weekly timetable of group work during one semester (14 wk). Each group is engaged in peer learning for 4 wk and then presents in a miniconference. Each vacant space in Figure 1 represents a week, during which students learn independently, using the course’s interactive website. The shaded weeks at the beginning and the end of the semester indicate when instructors lecture to the whole class. The four shaded weeks are the only time when teaching is in the form of lectures. Other than that, learning takes places either in teams of six students or in the larger group that consists of five teams.

Table 2. Example of group-learning topics and subtopics

Group topics	Team subtopics
Eukaryotic cell	Chloroplast Mitochondria Lysosome Golgi apparatus Eukaryotic cell complexity
Cellular membrane	Composition and structure Passive transport Active transport Endo- and exocytosis Signal transduction

Week	1	2	3	4	5	6	7	8	9	10	11	12	13	14
	Work on presentation				MiC									
		Work on presentation			MiC									
			Work on presentation		MiC									
				Work on presentation		MiC								
					Work on presentation		MiC							
						Work on presentation		MiC						
							Work on presentation		MiC					
								Work on presentation		MiC				
									Work on presentation		MiC			
										Work on presentation		MiC		
											Work on presentation		MiC	

Figure 1. Timetable for the group learning in the active-learning version.

Team Learning. The team members were expected to be engaged in a comprehensive study of the topic they chose and, as indicated, present it to the entire group. They had to look for relevant information, discuss their understanding with peers and the TA, and collaborate in designing their presentation and talk. The team learning was supported by a structured “team space” on the course website.

The team space in the course website included detailed guidelines for the consecutive working stages: individual learning of the topic, beginning to work, focusing on the team subtopic, preparing a draft presentation, revising the presentation, presenting to the TA and getting feedback, and presenting and discussing at the miniconference. A timeline was recommended for each stage, and a place for discussion and uploaded files was provided. In the second stage, “beginning to work,” we presented an experiment followed by questions to enhance curiosity and further learning of the subtopic. An example of such a team question is presented in Table 3.

TAs met their teams in the opening session of the larger group. They were guided to facilitate and mediate discussions rather than teach the content. During the following month, they were available almost daily on the team website, answering questions, commenting on team discussions and presentation drafts, and monitoring team progress. Sometime before their groups’ miniconferences, they met their teams to give final feedback on the presentations.

Miniconference. The summary and climax of the team learning was a miniconference for each “larger group” consisting of all the teams that learned the subtopics of one topic during the month (see examples of topics and subtopics in Table 2). In the miniconference, each team presented its subtopic to the other teams and the teaching staff. Each presentation was followed by a whole-group discussion led by the instructor, who added his input and asked challenging questions. He highlighted the connections between subtopics and provided information about relevant cutting-edge research as well.

Assessment

The traditional two assessment components of the course were a midterm quiz (worth 5% of the final score) and a final exam (worth 95% of the final score), both in the form of MC questions. The literature on meaningful learning acknowledges the limits of MC tests in developing deep learning. During our intervention, we added open-ended, higher-order thinking (O-HOT) questions to the final exam, thus testing higher-order thinking and reflecting the nature of instruction more sensitively (Linn *et al.*, 2006; Marx *et al.*, 2004). Open-ended questions require students to bring forth evidence and think like scientists by encouraging them to explore a variety of solutions (Schinske, 2011).

In the active-learning version, in addition to the exams, the TAs assessed each team’s learning process using a scoring rubric (see Supplemental Material), and the instructors assessed each team’s final presentation (see Table 1).

Procedure

Biology 1 is taught, in separate classes, to biology majors or students from affiliated programs, who enroll in the course in their very first semester, and to other STEM undergraduates, who are in at least their second semester. Altogether, during the time of the intervention, more than 2000 students enrolled to the course.

As previously mentioned, the actual change in the course was not identical to our initial design. This was because of the reluctance of the university to obligate students to participate in any “treatment group” and to cancel lectures if students wish to attend them. As we could not assign students randomly, the study was quasi-experimental.

In each semester, the students could choose their preferred version of the course. Because attendance was not mandatory, many students, especially nonbiology majors, preferred not to attend class at all. Students who did not attend class were identified as participating in the traditional-plus-tutorial version. The lectures students observed on the website were videotaped prior to the intervention, and they could also use all the other resources provided online. A student who

Table 3. An example of a driving question for the onset of team work

Protein degradation—Where have the mice brain proteins gone?

It is possible to track body materials *in vivo* by labeling them using radioactive isotopes. This approach is based on the fact that radioactive isotopes are processed as are natural substances (e.g., amino acids in proteins). The advantage of this method is that radiolabeled compounds can be detected using scintillation counters.

A researcher fed mice with [¹⁴C]-labeled lysine (note that the half-life time ¹⁴C is 5770 yr). He expected that after some time most mice proteins would contain the radiolabeled lysine and that the level of radioactivity in mice proteins would be proportional to the amount of [¹⁴C]lysine in tissues. The mice mated and reproduced. Females kept on eating radioactive food throughout pregnancy (~3 wk) and nursing (~3 wk). Their offspring ate [¹⁴C]lysine until maturity (~60 d), at which time all body proteins became radioactive. When the offspring were 60-d old, radioactive food was replaced by normal food. Once radioactive food was not available, mice started to synthesize radioactive-free proteins.

Subsequently, at time intervals after shifting to normal food, mice were sacrificed, and the level of radioactivity in brain tissues was monitored. Results are illustrated in the following graph:



When preparing your team presentation, please refer to the following questions:

1. How can you explain the results of the experiment?
2. Why is it reasonable to assume that protein degradation did not happen in lysosomes?
3. What is the link between ubiquitin, a protein discovered by Technion Nobel laureates Hershko and Ciechanover, and the phenomenon described in the graph?

A question you should consider, but not necessarily include in your presentation: Is it possible that the observed decrease in radioactivity was the result of natural degradation of ¹⁴C?

indicated he/she attended most classes when the adapted teaching was carried out was tagged as participating in that version, as were students who enrolled in the active-learning version. Only in the last semester did the university allow lectures to be canceled, which enabled students to choose either the active-learning group or to settle for the individual learning (traditional plus tutorial). The students who enrolled in the active-learning group reported on their motives. The main motive was hope for better course grade, which is not based only on the final exam in the active-learning group. Some students believed this version would force them

to learn throughout the semester, rather than for the exam only, and a few indicated learning difficulties and preferred close contact with a TA. Table 4 shows the various course versions across the 3-yr research, and a data collection sample that will be explained in the following section. As indicated, although we originally planned the course versions to be implemented in consecutive versions, the university constraint eventually became an advantage in the data collection design, since having two or three versions during one semester allowed us to have comparison groups from the same class.

Table 4. Course versions and data collection sample for statistical analysis (*n* = 569)

Approach	2009		2010		2011	
	Spring ^a	Winter ^a	Spring	Winter	Spring	Winter
Traditional plus tutorial	—	—	61	77	171	—
Adapted teaching	—	—	44	162	—	—
Active learning	—	—	—	28	—	26

^aThe consecutive semesters Spring 2009 and Winter 2010 are presented only to indicate the teaching version, although we do not present student data from these semesters.

Data Collection

Data collection included the final achievement tests and class observations in the form of videotapes and journal entries made by M.T. In the year prior to the study (2008), we videotaped all the lectures to follow teaching and learning patterns before the intervention and to have the videos for use in the tutorial developed by O.S. (Sagy *et al.*, 2011).

A major reflection of the course instructor that directed our data collection was that during the implementation of the adapted-teaching version students asked more thoughtful questions in class compared with previous years. Consequently, we scrutinized all student questions from the videotapes (made in Spring 2008, before intervention) and from the researchers' journal (in the adapted-teaching version, Spring 2010). To compare the two, we selected two sessions per semester that dealt with the same topic and spanned the same time. These class sessions (lecture periods) dealt with 1) membranes (90 min) and 2) differentiation (45 min) and were taught by the same instructor in both years. In both semesters, the course was taught in large auditoriums, and the number of students who attended class was similar (about 40–50).

Student performance data were collected from Spring 2008 (prior to intervention) and during 2009–2011. Despite the overall number of students who enrolled in the course during the 3 yr, the comparative data sample we present here is smaller ($n = 569$; see Table 4), as explained later in this section.

We compared the achievements of students by their learning patterns. As indicated, students could choose their preferred learning pattern (i.e., attending/not attending class, which determined whether they were associated with the traditional-plus-tutorial or with the adapted-teaching versions or enrolling in the active-learning version).

To ensure credible comparison, we included in the sample presented in Table 4 only students who: 1) studied during the same semester and took the same exam; 2) took the first term exam¹; 3) were classified according to the learning pattern they reported in a voluntary self-reported questionnaire (attended vs. did not attend class; individual vs. group learning); and 4) took the (voluntary) precourse test.

To avoid possible bias resulting from students' unequal assignment to "treatments" with respect to prior knowledge, a pretest was administered prior to the introduction of the active-learning version. The pretest consisted of questions from the matriculation exams for high school biology majors, relevant to topics studied in the Biology 1 course. No significant difference was found in pretest scores between the students' prior knowledge in the different groups.

Data Analysis

Students' Questions. Based on Anderson and Krathwohl (2000) and in line with Shepardson and Pizzini (1991) and Marx *et al.* (2004), students' questions in class were classified into three cognitive levels. Borrowing Shepardson and

Pizzini's terminology, input (low-level) questions addressed merely factual knowledge, for example: "Is Cdk a protein?" The processing (medium) level required students to draw relationships within and between information, data, and principles. The students needed to compare, contrast, and apply knowledge, for example: "Does the protein regulating the last stage of the cell cycle act by the same mechanism as the protein regulating the first stage?" The output (higher) level questions required students to go beyond the data at hand and use them to hypothesize, generalize, or predict, for example: "If we administer a drug that stops the cell cycle, how can we prevent it stopping the cycle in all our body cells?" The primary classification was done by M.T., and then was further confirmed by T.T, O.S., and another researcher.

The Final Exam. Traditionally, the final exam for the course consisted of only MC items. In light of the aforementioned literature, which advocates the use of open-ended questions that allow higher complexity, we added to the final exam open-ended questions dealing with either content or scientific processes (Marx *et al.*, 2004). These questions required four types of thinking skills that we believe students should have acquired during the course: 1) articulating a biological principle drawn from the given data (Ramsden, 1992); 2) describing a mechanism instead of supplying "black box" answers (Chin and Brown, 2000); 3) using a correct, evidence-based argument (Kuhn, 1993); and 4) doing near transfer, that is, using knowledge in a different context within the course syllabus (Sasson and Dori, 2012). Table 5 presents an open-ended question and the skills required to respond to each part.

For scoring these open-ended items, we developed a rubric that was approved by the two course instructors, D.Z. and S.G., who are biology professors, and two researchers who hold PhDs in biology education. The scoring of student responses according to this rubric was done by M.T., who is an experienced high school biology teacher with experience in teaching undergraduate students as well. Statistical analysis was performed on student achievement test scores. Mean comparisons were done using analysis of variance, and then multiple comparisons were applied to the pairwise means comparisons. Their effect with respect to size was evaluated using effect-size correlation r . Interpretative analysis was carried out on observational data.

RESULTS

With the change the nature of the lectures, the adapted-teaching version became the first step in diverting the responsibility for learning informative content to the students. The instructors' attempts to refrain from merely delivering content enabled them to allocate more time to promoting understanding of biological research processes and the connections between basic and applied science. The instructors' reports of a greater quantity of "better questions" asked by the students led to our attempt to examine the questions consistently.

Students' Questions

As indicated, we selected two class sessions: one from the year prior to the intervention (2008) and another from the year of the adapted-teaching version (2010) to compare

¹In Israel, students are allowed to take each final exam twice. They can take the second term (make-up exam) instead of the first, because of scheduling constraints or to improve their score. In cases in which they take the second term exam in addition to the first, the second exam score will be their final score.

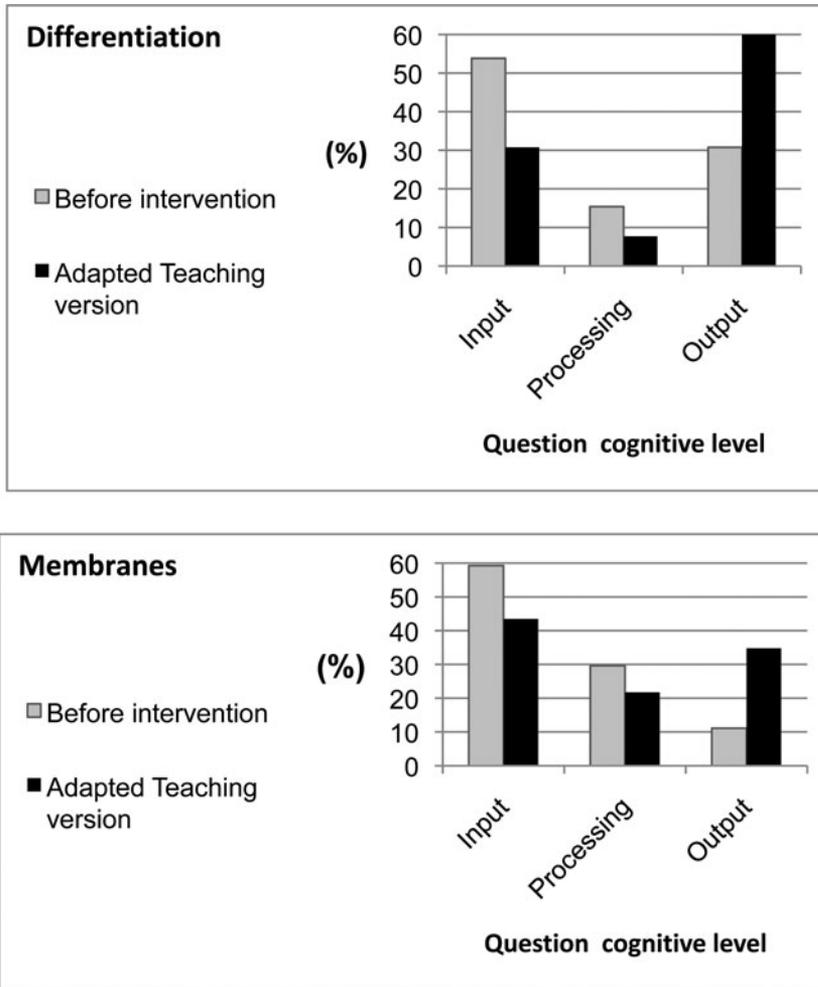


Figure 2. Student questions of various cognitive levels in two topics asked in the adapted-teaching version (Spring 2010) and before intervention (Spring 2008).

required higher-order thinking (eight out of 40 MC questions and three open-ended).

Student Performance in the Active-Learning Version

After establishing the foundation for independent learning, the following step was to further improve the depth of learn-

ing by a substantial decrease in the number of lectures, in which students are merely passive learners. Instead, students were expected to study the majority of the topics independently with the support of the online tutorial and take part in active small-group investigations of one topic.

Although we meant to cancel most lectures to allow the instructors to function as mediators of group discussions, the

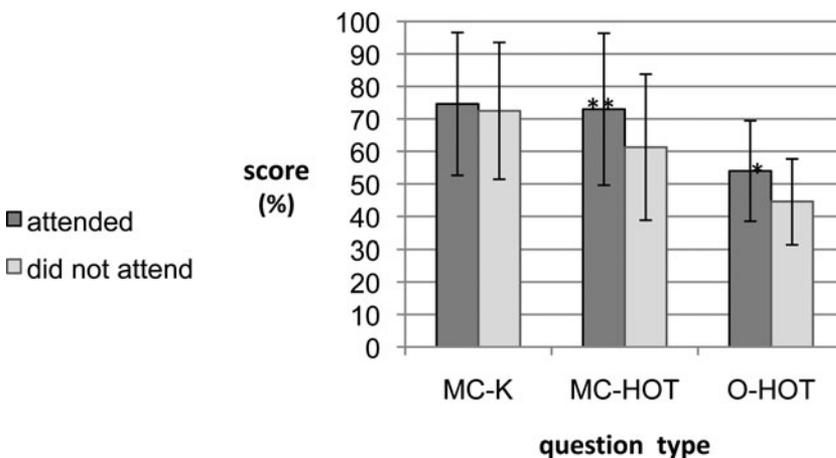


Figure 3. Achievements of students studying in the adapted-teaching version (attended class) vs. these studying in the traditional-plus-tutorial version (did not attend class) in Spring 2010.

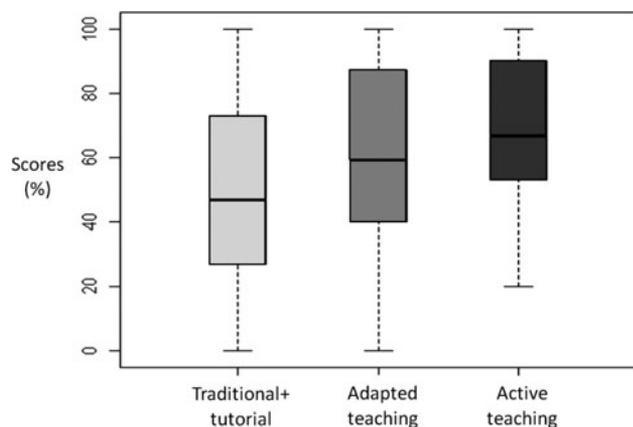


Figure 4. Student performance in the open-ended questions across the three versions of the course (Winter 2011).

university did not originally permit canceling lectures for those students who wish to attend class. Therefore, in the Winter semester of 2011, students could, in fact, take part in each of the three versions of the course (“treatments,” see Table 4). Those preferring only individual learning represented the traditional-plus-tutorial version; students who attended class represented the adapted-teaching version; and those who in addition to individual learning took part in group learning of one topic represented the active-learning version. The comparison of the three groups is presented in Figure 4. As indicated, pretest scores of those students show no significant difference. As in the previous stage, we found no difference among the three groups with respect to MC items, which mainly require memorization. Most interesting is what we found with respect to the most complex, open-ended items.

As can be seen in Figure 4, O-HOT scores for questions in the active-learning version group (experimental) were higher than those of the adapted-teaching version and the traditional-plus-tutorial version comparison groups ($p < 0.09$ and 0.004 ; effect size = 0.18 and 0.33 , respectively).

In the following semester (Spring 2011), most lectures were canceled, allowing another comparison between student performance in the traditional-plus-tutorial version and the active-learning version. In that semester, a group of 26 students enrolled in the experimental 1-mo active learning, while the others made up the comparison group. Table 6 presents student scores in the open-ended items in the active-learning group and the traditional-plus-tutorial group across two consecutive semesters. Here again, we found no significant difference between the groups in the scores of the MC items. Although all the O-HOT questions dealt with topics that were not taught to the experimental 1-mo learning groups, the

students who enrolled in the active-learning group outperformed their counterparts who preferred only independent learning.

We compared student achievements only within semesters to avoid interfering variables: different instructors, different students in different semesters, different exams. As can be seen in Table 6, gaps between semesters are greater than between treatments; however, we cannot address gaps between semesters, for there are too many variables involved.

DISCUSSION

The original course pedagogy, of lectures given in a large hall, reflects the past requirements, as is evidenced in the final test. This test was based mainly on MC items that required memorization of content. Across all treatments in the three versions, the results of the overall MC part of the final exam were similar. Students who learned independently, using the online tutorial (which included the lectures), had scores comparable with those of the students who participated in the other two treatments (adapted teaching and active learning). Thus, if learning biology means doing well on an MC knowledge test, then the university and the faculty of biology can keep lecturing, as this form of teaching is the most effective in terms of cost. Nevertheless, if one is expecting more than the ability to recall knowledge, then other forms of teaching and assessment should be considered. In our study of the Biology 1 course, we were interested in more sophisticated learning and in methods that enhance and reflect students’ deep learning.

The adapted-teaching version was the first step in reducing the instructor’s responsibility for covering all content and increasing students’ responsibility for learning, using the course website as a supportive active-learning tool. In this version, active construction of conceptual knowledge was enhanced, since some of the content was independently learned. Moreover, the instructor could make references to broader issues to a greater extent than before (for example, the Nobel Prize of that year, awarded for discovering the ribosome structure), or he could expand complex or difficult topics raised in students’ questions. Such adaptations to the class level and the context without changing the syllabus were discussed by Davis and Varma (2008), who emphasized the importance of designing a curriculum in a way that enables instructors to make adaptations. In line with Handelsman *et al.*’s (2004) characterization of scientific teaching, such adapted teaching offers more understanding of scientific research processes and exposes students to the limitations of science, as well as to its power. Evidence for how the change in instruction in the adapted-teaching version promoted deep learning was expressed by the difference in the nature of students’ questions during the lectures, and by their superior

Table 6. Student O-HOT questions scores in the active-learning (experimental) version vs. the traditional-plus-tutorial (comparison) version

Version	Active learning		Traditional plus tutorial		<i>t</i> -test value	<i>p</i>	Effect size
	Mean score (SD)	<i>n</i>	Mean score (SD)	<i>n</i>			
Winter 2011	34.6 (22.6)	28	23.5 (21.7)	42	2.05	0.04	0.24
Spring 2011	61.6 (24.3)	26	50.4 (25.8)	171	2.08	0.04	0.22

performance in the higher-order thinking parts of the final exam compared with the performance of students who learned independently.

In the active-learning version, social-constructivist learning was promoted by three factors: 1) peer learning, in which students shared and negotiated knowledge with their peers while producing artifacts (the PowerPoint presentation); 2) TA tutoring, which was encouraging and challenging, as well as reflective (Wood and Tanner, 2012); and 3) the miniconference-like lesson, in which students presented their understanding of a topic to a larger group and the instructor and received feedback. At that stage, we found that student outcomes in the open-ended questions of the final exam were higher than those of the comparison groups. It has already been shown that short peer discussions enhance understanding and consequently improve achievements in concept questions (Smith *et al.*, 2009). We strongly believe that the group learning during the 1 mo in which the students negotiated their understanding with peers and with the TAs and worked together to produce an artifact of their learning (the final presentation) contributed a great deal to their deep learning. We assert that group learning during 1 mo developed a “biological discourse” as well, which is more typical of advanced stages of learning biology. Formative assessment given by peers and the TA while working together on the presentations further developed the students’ learning outcomes. We suggest that the final miniconference, in which each team presented its knowledge of a topic and was subject to summative assessment (by the instructor), had an additional effect on the students’ outcomes in higher-order thinking assignments in the final exam.

In-depth coverage of a topic can elaborate incomplete ideas, provide different views of a phenomenon, enhance abstract thinking, and elicit abstract reasoning (Eylon and Linn, 1988). All these happened during the 1-mo group study while the students constructed their knowledge through comprehensive independent learning and maintained a continuous discourse in the team and with the TA and the instructor. We further assume that this 1-mo group-learning experience enhanced patterns of deep learning of the entire course content. This can be seen from the performance in the open-ended items that did not address topics discussed in the groups. The students apparently applied to the rest of the course topics those biological principles, reasoning patterns, and ways of thinking and learning they had developed while learning “their own” topic. The scope of this paper does not allow us to present qualitative data from students’ interviews, but based on the aforementioned data, we argue that students who experienced the team and group learning have developed patterns of deep learning across topics.

An additional component of the instructional model that promoted deep learning was an assessment process that no longer required mere rote learning, because it included the additional open-ended items in the test. The formative and summative assessment of the team learning encouraged deep learning as well. This is congruent with the idea that assessment has an impact on the nature of learning (Marton and Saljo, 1976; Black *et al.*, 2003). The working guidelines posted on the team website, and especially the ongoing dialogue

with the TA, made it clear to the students that they were expected to go beyond rote learning, although this part of the course focused on only one topic. In this regard, Biggs and Tang (2007) found that assignments focused on one topic had a positive effect on the nature of learning, while assignments that required coverage of range of topics encouraged students to adopt a surface approach.

Our findings also add to what Schwartz *et al.* (2009) found regarding the importance of in-depth learning of one topic. They studied student performance in biology, chemistry, and physics higher education introductory courses with respect to the nature of learning these subjects in the high school. They found that in-depth learning of at least one major scientific topic in high school for a month or more influenced performance in college more positively than covering all major topics in breadth.

Thus, we suggest that at the undergraduate level, studying even one topic in depth makes a difference and causes learners to adopt a deep-learning approach. In biology, dealing with complex representations is common, even in an introductory course. Understanding complexity requires multiple opportunities be given to students for constructing their understanding. Moreover, multiple experiences, in which ideas and phenomena may be coordinated into richer and more complex understanding, are needed to establish conceptual understanding. This process is intensive and time-consuming (Schwartz *et al.*, 2009). The 1-mo group learning that enhanced our students’ deep learning of one topic offers such opportunities. However, the acquired learning habits and deep understanding constructed in this intensive and multifaceted learning experience were apparently transferred to the learning of other course topics.

From a practical point of view, taking into consideration the current constraints of the university, the instructional model we implemented is actually sustainable. After the online tutorial is developed and established, it only needs maintenance and updating. The scope of this paper does not allow us to go into the process the TAs went through nor to describe how they developed as mentors. However, Dolan and Johnson (2009) have already pointed to the possible merits of employing TAs. They suggest that the mentorship experience could help graduate students (scientists-in-training) to improve their understanding of scientific issues and acquire communication skills, while developing their identities as scientists.

Finally, we believe that the contribution of the active-learning instructional model is twofold and goes beyond merely improving student performance. First, this model, which implemented a social-constructivist approach in which learning took place through discourse with peers and staff and through the carefully designed website, exposed students to the essence of biology as a research discipline and to the ways biologists interpret phenomena and investigate them. Second, we hope that the students’ perception of the professor as a researcher and a role model replaced the perception of the professor as a mere deliverer of scientific content. We attach great importance to both in the future development of all students: those who major in life sciences and those others who get to know biology as engineers or scientists in other fields.

REFERENCES

- Anderson LW, Krathwohl DR (eds.) (2000). *A Taxonomy for Learning, Teaching and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives, Complete ed.*, New York: Longman.
- Anderson W *et al.* (2011). Changing the culture of science education at research universities. *Science* 331, 152–153.
- Ash D. (2004). Reflective scientific sense-making dialogue in two languages: the science in the dialogue and the dialogue in the science. *Sci Educ* 88, 855–884.
- Biggs J, Tang C (2007). *Teaching for Quality Learning at University*, 3rd ed., Maidenhead, UK: Open University Press/McGraw-Hill Education, 195–216.
- Black P, Harrison C, Lee C (2003). *Assessment for Learning: Putting It into Practice*, New York: Open University Press.
- Chin C, Brown DE (2000). Learning in science: a comparison of deep and surface approaches. *J Res Sci Teach* 37, 109–138.
- Davis EA, Varma K (2008). Supporting teachers in productive adaptation. In: *Designing Coherent Science Education: Implications for Curriculum, Instruction, and Policy*, ed. Y Kali, M Linn, and JE Roseman, New York: Teachers College Press, 94–122.
- Dolan E, Johnson D (2009). Toward a holistic view of undergraduate research experiences: an exploratory study of impact on graduate/postdoctoral mentors. *J Sci Educ Technol* 18, 487–500.
- Dori YJ, Herscovitz O (1999). Question-posing capability as an alternative evaluation method: analysis of an environmental case study. *J Res Sci Teach* 36, 411–430.
- Driver R, Asoko H, Leach J, Scott P, Mortimer E (1994). Constructing scientific knowledge in the classroom. *Edu. Res* 23, 5.
- Ennis RH 1987. A taxonomy of critical thinking dispositions and abilities. In: *Teaching for Thinking*, ed JB Baron and RJ Sternberg, New York: WH Freeman, 9–26.
- Eylon BS, Linn MC (1988). Learning and instruction: an examination of four research perspectives in science education. *Rev Educ Res* 58, 251–301.
- Gilmer P (2010). *Transforming University Biochemistry Teaching Using Collaborative Learning and Technology: Ready, Set, Action Research!*, Dordrecht, Netherlands: Springer.
- Haak DC, HilleRisLambers J, Pitre E, Freeman S (2011). Increased structure and active learning reduce the achievement gap in introductory biology. *Science* 332, 1213–1216.
- Handelsman J *et al.* (2004). Scientific teaching. *Science* 304, 521–522.
- Henderson C, Beach A, Finkelstein N (2011). Facilitating change in undergraduate STEM instructional practices: an analytic review of the literature. *J Res Sci Teach* 48, 952–984.
- Hofstein A, Navon O, Kipnis M, Mamlok-Naaman R (2005). Developing students' ability to ask more and better questions resulting from inquiry-type chemistry laboratories. *J Res Sci Teach* 42, 791–806.
- Klymkowsky MW, Garvin-Doxas K, Zeilik M (2003). Bioliteracy and teaching efficacy: what biologists can learn from physicists. *CBE Life Sci Educ* 2, 155.
- Kuhn D (1993). Science as argument: implications for teaching and learning scientific thinking. *Sci Educ* 77, 319–337.
- Linn MC, Hsi S (2000). *Computers, Teachers, Peers: Science Learning Partners*, Mahwah, NJ: Lawrence Erlbaum.
- Linn MC, Lee HS, Tinker R, Husic F, Chiu JL (2006). Teaching and assessing knowledge integration in science. *Science* 313, 1049.
- Marbach-Ad G, Sokolove PG (2000). Can undergraduate biology students learn to ask higher level questions? *J Res Sci Teach* 37, 854–870.
- Marton F, Saljo R (1976). On qualitative differences in learning—II: outcome as a function of the learner's conception of the task. *Br J Educ Psychol* 46, 115–127.
- Marx RW *et al.* (2004). Inquiry-based science in the middle grades: assessment of learning in urban systemic reform. *J Res Sci Teach* 41, 1063–1080.
- Pfund C, *et al.* (2009). Summer institute to improve university science teaching. *Science* 324, 470.
- Ramsden P (1992). *Learning to Teach in Higher Education*, London: Routledge.
- Sagy O, Kali Y, Zilberstein D, Tsaushu M, Tal T, Gepstein S (2011). To Learn or to Be Taught? Harnessing Technology to Enhance Self Regulated Learning, Raanana: The Open University of Israel.
- Sasson I, Dori YJ (2012). Transfer skills and their case-based assessment. In: *The Second International Handbook of Science Education*, ed. BJ Fraser, KG Tobin, and CJ McRobbie, Dordrecht, Netherlands: Springer, 219–250.
- Schinske JN (2011). Taming the testing/grading cycle in lecture classes centered around open-ended assessment. *J Coll Sci Teach* 40, 46–52.
- Schwartz MS, Sadler PM, Sonnert G, Tai RH (2009). Depth versus breadth: how content coverage in high school science courses relates to later success in college science coursework. *Sci Educ* 93, 798–826.
- Scouller K (1998). The influence of assessment method on students' learning approaches: multiple choice question examination versus assignment essay. *Higher Educ* 35, 453–472.
- Scouller KM, Prosser M (1994). Students' experiences in studying for multiple choice question examinations. *Stud Higher Educ* 19, 267–279.
- Shepardson DP, Pizzini EL (1991). Questioning levels of junior high school science textbooks and their implications for learning textual information. *Sci Educ* 75, 673–682.
- Smith MK *et al.* (2009). Why peer discussion improves student performance on in-class concept questions. *Science* 323, 122–124.
- Wood WB, Tanner KD (2012). The role of the lecturer as tutor: doing what effective tutors do in a large lecture class. *CBE Life Sci Educ* 11, 3–9.
- Woodin T, Carter VC, Fletcher L (2010). Vision and Change in Biology Undergraduate Education, A Call for Action—initial responses. *CBE Life Sci Educ* 9, 71–73.
- Woodin T, Smith D, Allen D (2009). Transforming undergraduate biology education for all students: an action plan for the twenty-first century. *CBE Life Sci Educ* 8, 271–273.