

Article

Using Targeted Active-Learning Exercises and Diagnostic Question Clusters to Improve Students' Understanding of Carbon Cycling in Ecosystems

April Cordero Maskiewicz,* Heather Peckham Griscom,[†] and Nicole Turrill Welch[‡]

*Point Loma Nazarene University, San Diego, CA 92106; [†]James Madison University, Harrisonburg, VA 22807;

[‡]Mississippi University for Women, Columbus, MS 39701

Submitted February 12, 2011; Revised November 3, 2011; Accepted November 21, 2011

Monitoring Editor: Mary Pat Wenderoth

In this study, we used targeted active-learning activities to help students improve their ways of reasoning about carbon flow in ecosystems. The results of a validated ecology conceptual inventory (diagnostic question clusters [DQCs]) provided us with information about students' understanding of and reasoning about transformation of inorganic and organic carbon-containing compounds in biological systems. These results helped us identify specific active-learning exercises that would be responsive to students' existing knowledge. The effects of the active-learning interventions were then examined through analysis of students' pre- and postinstruction responses on the DQCs. The biology and non-biology majors participating in this study attended a range of institutions and the instructors varied in their use of active learning; one lecture-only comparison class was included. Changes in pre- to postinstruction scores on the DQCs showed that an instructor's teaching method had a highly significant effect on student reasoning following course instruction, especially for questions pertaining to cellular-level, carbon-transforming processes. We conclude that using targeted in-class activities had a beneficial effect on student learning regardless of major or class size, and argue that using diagnostic questions to identify effective learning activities is a valuable strategy for promoting learning, as gains from lecture-only classes were minimal.

INTRODUCTION

The national call to transform undergraduate biology education includes the challenge for instructors to help students develop scientific ways of thinking (National Research Council [NRC], 2003; American Association for the Advancement of Science [AAAS], 2010). An essential aspect of thinking like a scientist, particularly a biologist, involves applying fundamental scientific principles to reason about and develop explanations of the living world. Yet we often find that stu-

dents rely on informal experiences and reasoning when trying to make sense of or explain biological systems (Songer and Mintzes, 1994; Carlsson, 2002a,b; Treagust *et al.*, 2002; Maskiewicz, 2006; Wilson *et al.*, 2006; Mohan *et al.*, 2009; Hartley *et al.*, 2011).

Consider, for example, the following synthesis question and one college-level biology major's response. The question was published by Ebert-May *et al.* (2003) and is now part of a larger set of diagnostic questions designed to elicit students' conceptions and ways of reasoning about biological and ecological phenomena related to transformations of matter and energy.

Question: Grandma Johnson requested to be buried in a canyon. Describe below the path of a carbon atom from Grandma Johnson's remains, to inside the leg muscle of a coyote. Be as detailed as you can be about the various molecular forms that the carbon atom might be in as it travels from Grandma Johnson to the coyote. NOTE: The coyote does not dig up and consume any part of Grandma Johnson's remains.¹

DOI: 10.1187/cbe.11-02-0011

Address correspondence to: April Cordero Maskiewicz (AprilMaskiewicz@pointloma.edu).

© 2012 A. C. Maskiewicz *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.

¹Adapted from Ebert-May *et al.*, 2003.

Student response: Carbon decomposes and feeds a bush with nutrients . . . The bush turns the carbon into energy and then excretes it into the atmosphere in the form of carbon dioxide. Coyote inhales and the oxygen populates the bloodstream giving cells this ability to carry nutrients to the muscle.

This student appears to be relying on everyday reasoning at an organismal level only to account for the path of a carbon atom, instead of applying a scientific, multilevel understanding of the transformation of matter. For example, the student states that carbon, rather than Grandma Johnson, decomposes and adds that “nutrients” moved to the plant, and then confuses the gas exchange between plants and animals. Underlying these unscientific conceptions is the fact that neither matter nor energy is conserved in this response; the student believes carbon “turns into” energy or changes into oxygen. Most preassessment responses to the “Grandma Johnson” question from this current study and others (Hartley *et al.*, 2011) reveal that biology majors, both upper and lower division, tend to apply such everyday reasoning when asked to trace the path of a carbon atom through a food chain, instead of providing explanations constrained by the fundamental principles of matter conversion. It follows that an appropriate instructional goal would be to help students develop ways of thinking about the relationship between scientific principles and transformative biological phenomena. In addressing that goal, this paper describes the results of three college-level biology instructors’ use of targeted active-learning activities and diagnostic assessments in upper- and lower-division ecology courses and contrasts student learning in those classes with results from a more traditional, lecture-based ecology course.

Background: The DQC Project

The impetus for this research emerged as a result of the authors’ 3-yr participation in the diagnostic question cluster (DQC) faculty development program (Hartley *et al.*, 2011). This program integrates education research with professional development (PD) in an effort to help faculty use DQCs and targeted active-learning activities to improve student reasoning and understanding about how matter and energy move within and through biological systems. DQCs are sets of interrelated questions about core biological concepts and ideas designed to identify problematic patterns in students’ thinking about biological content (D’Avanzo, 2008). The questions are effective in helping instructors deconstruct their students’ understanding to reveal small-scale problems that limit large-scale understanding. The DQCs used in this study are informed by an overarching framework based on three principles: conservation of matter, conservation of energy, and scales of organization (see *Thinking Like a Biologist* website [www.biodqc.org] for framework and assessment questions). Engaging faculty in the scholarship of teaching and learning is a secondary objective of the DQC program, and it is as participants in this program that we report here on the outcomes of pairing DQCs with targeted active-learning activities.

Targeted Active-Learning Exercises and DQCs

In our study, we used the results of the diagnostic questions to guide the development and implementation of active-

learning exercises specifically targeted to respond to students’ current reasoning and understanding within a particular topic. Designing instruction that builds on students’ existing knowledge emerged from work by Piaget (1963/2003) and has since been modified to incorporate the idea that students construct their own knowledge within social and cultural activities (Brown *et al.*, 1989). Student-centered approaches to learning emerged from this research as “active learning,” and the education literature has established that active learning improves student understanding more than traditional, passive lectures (Hake, 1998; NRC, 2000, 2003; Beichner and Saul, 2003; AAAS, 2009, 2010; Minner *et al.*, 2010). Active-learning environments encourage students to work together solving problems, exploring relationships, analyzing data, and articulating and comparing points of view (Johnson *et al.*, 1998). Within biology, much of the active-learning research examines students’ learning over an entire course, and the findings demonstrate the effectiveness of active learning in increasing academic performance in both large lecture classes (Smith *et al.*, 2005; Armbruster *et al.*, 2009) and small classes (Michael, 2006).

Based on our knowledge of how people learn, active-learning exercises should be responsive to students’ existing knowledge; yet AAAS (2009) recommends that biology faculty decide, prior to instruction, what they want their students to know or be able to do so that learning goals can be determined. With conceptual inventories established as valid and reliable measures for assessing and diagnosing students’ knowledge and reasoning (Hake, 1998; NRC, 2006; D’Avanzo, 2008), we propose that these easy-to-use diagnostic test results can be analyzed prior to instruction to help inform learning outcomes based on a particular group of students’ existing knowledge. As such, these outcomes “can serve as a guide for selecting teaching strategies that will engage students and help them advance their understanding to the desired level of comprehension” (AAAS, 2009, p. 7).

Research shows that combining diagnostic questions with active-teaching approaches has the potential to improve students’ understanding and reasoning (Hestenes, 1998; Nehm and Reilly, 2007) and provides a validated way to measure the effectiveness of activities targeted at specific learning objectives. In this study, we used our DQC results to inform our instructional design, which allowed us to concentrate on how to advance learning and reasoning about the specific concepts of matter and energy identified by the literature as particularly difficult for students: transformation of inorganic and organic carbon-containing compounds in biological systems. By analyzing students’ DQC responses, the instructors were informed about their students’ knowledge, and as such, could better frame the content in ways that helped students develop a more robust understanding of biology. Thus, our findings contribute to identifying the elements of active learning most effective for different groups of students and specific topics.

Why Matter and Energy?

The topic of matter and energy transformations and pathways is one of the five core conceptual areas in biology identified in the recently published *Vision and Change: A Call to Action* (AAAS, 2010) report. The conservation of matter and energy are two central principles that biologists apply when reasoning about dynamic systems in which matter and

energy are exchanged across defined boundaries (NRC, 2003; AAAS, 2010). Because matter and energy cannot be created or destroyed, this principle can guide the development of an explanation by following inputs and outputs within a biological process or system. Consider, for example, the “Grandma Johnson” question at the opening of this paper. A scientific response to this question would account for the matter by tracing carbon from decomposers to the atmosphere via oxidation processes, then to a plant via photosynthesis, and finally to an herbivore and then a coyote via digestion and biosynthesis. Detailed understandings of these processes are not necessary, but applying principles of conservation can help the student reason through the problem scientifically. Therefore, we adopt the term “principle-based reasoning” to mean applying the scientific principles of nature to constrain one’s biological explanation (Mohan *et al.*, 2009; Hartley *et al.*, 2011).

As previously stated, research shows that undergraduate students tend to rely on informal reasoning when developing explanations of biological phenomena (i.e., making inappropriate inferences based on personal experiences or commonly used informal language; Carlsson, 2002a,b; Treagust *et al.*, 2002; Maskiewicz, 2006; Wilson *et al.*, 2006; Mohan *et al.*, 2009; Hartley *et al.*, 2011). Analysis of students’ responses to a set of associated DQC questions can reveal these types of patterns in students’ reasoning. For instance, if a student responds to the “Grandma Johnson” question by stating that a bush turns carbon into energy, and then later chooses a multiple-choice response that states, “Once carbon enters a plant, it can be converted to energy for plant growth,” then one can begin to conclude that this student is not applying the principle of conservation of matter when reasoning about plant growth. This example illustrates that how one reasons influences the explanations one develops; thus, knowledge and reasoning are fundamentally linked. By asking application questions, as opposed to questions on the details of biological processes, the results from the ecology DQCs illuminate reasoning patterns that are consistent for a student or even for an entire class.

The purpose of this study, then, was to implement targeted active-learning activities that address confusion with matter and energy and subsequently examine the effects on student understanding and reasoning by comparing pre- and post-DQC responses. Because the DQCs are specifically designed to diagnose how one traces energy and matter through three levels of biological complexity (atomic/molecular/cellular, organismal, and ecosystem), the ways one applies—or does not apply—the principles of conservation of matter and energy are revealed in students’ responses. Students applying principle-based reasoning are able to trace matter and energy across scales without thinking that matter can disappear or become energy in a biological context. Many studies show that students at both the secondary and college level have difficulty tracing matter and energy in biological and physical systems (Anderson *et al.*, 1990; Driver *et al.*, 1994; Wilson *et al.*, 2006; Mohan *et al.*, 2009; Hartley *et al.*, 2011), therefore, activities that help students develop the practice of tracing matter and energy as it is transformed during biological processes can facilitate students’ ability to make connections between different levels of organization in biology and, thus, begin to account for the complexities of biological systems. Furthermore, this specific focus on only matter and energy allowed

us to uncover nuances in student knowledge *and* reasoning about matter and energy at multiple scales and across multiple contexts. Such knowledge provides insight into additional instructional interventions that can help students progress in their development of principle-based reasoning.

METHODS

In this quasi-experimental study, we examine the effects of the active-learning interventions on students’ reasoning about carbon transformations through both quantitative and qualitative analysis of students’ pre- and postinstruction responses to topic-specific DQCs. The combination of both quantitative and qualitative analysis of students’ responses is necessary for identifying the reasoning resources students already have that can be built upon, as well as revealing changes in students’ reasoning postinstruction. We also include a brief discussion of how specific student-centered activities were chosen based on findings from pretest results.

Settings and Subjects

The four faculty who contributed data to this study were self-motivated and interested in using targeted active-learning exercises and assessment tools in their classrooms to improve upon students’ initial reasoning. Each volunteered for the PD program supported by a National Science Foundation Division of Undergraduate Education (NSF-DUE) grant focusing on the use of targeted active-learning activities and DQCs. The DQC PD program involved faculty from a broad range of institutions and incorporated multiple components, including: 1-d workshops at the 2008, 2009, and 2010 annual meetings of the Ecological Society of America; faculty collaborations; dissemination of findings and experiences; and interaction among faculty and program leaders via email and telephone regarding coding of answers, use of active-learning strategies, and help with data analysis.

Similar to the participants in the DQC program itself, the instructors contributing to this study come from a broad range of institutions as defined by the Carnegie Foundation Classifications (<http://classifications.carnegiefoundation.org>). Instructor 1 (author) teaches at a private, not-for-profit, 4-yr, selective university of approximately 2400 undergraduates. Instructor 2 (author) teaches at a public, 4-yr, more-selective school of 18,000 undergraduates. Instructor 3 (author) teaches at a public, 4-yr, inclusive university of approximately 2500 undergraduates. Instructor 4 (not an author of this paper) teaches at a private, not-for-profit, 4-yr, selective school of approximately 4600 students.

Instructional Intervention

The majority of students participating in this study were biology majors, with the exception of Instructor 1’s students (Table 1). Varied use of active-learning by the four instructors permits us to comment on both the effectiveness of active versus passive instructional methods, as well as faculty use of paired-activity diagnostic strategies. The percentage of class time spent on active learning was determined based on review of syllabi and reflection notes recorded on daily lesson plans during units of instruction that discussed matter

Table 1. Characteristics of the study participants and their institutions

Instructor	1-AL	2-AL	3-MAL	4-L
Instructional method ^a	86% active learning	80% active learning	40% active learning	<5% active learning
Student population	Non-science majors: freshman to senior	Biology majors: lower division	Biology majors: upper division	Biology majors: lower division
Course	General education: Ecology and Conservation	Core course: Ecology and Evolution	Upper division: General Ecology	Core course: Ecology
Number of students in course ^b	42	70	19	70
University description	Private, not-for-profit, 4-yr, selective university of approximately 2400 undergraduates	Public, 4-yr, more-selective school of 18,000 undergraduates	Public, 4-yr, inclusive university of approximately 2500 undergraduates	Private, not-for-profit, selective, 4-yr school of approximately 4600 students
University average SAT	1129	1146	1040	1054
University demographic (% female)	69	60	82	63

^aRemaining percentage of instructional time was passive lecture format. An instructor using active learning 80% of the time spends 20% of her teaching time using more passive delivery methods.

^b"n" in all statistical analyses.

and energy. Instructors 1 (Maskiewicz, identified hereafter as "Instructor 1-AL" ["AL" indicating active learning was used throughout the course]) and 2 (Griscom, identified hereafter as "Instructor 2-AL") spent 86 and 80% of class time, respectively, on active learning, while Instructor 3 (Welch, hereafter referred to as "Instructor 3-MAL" ["MAL" referring to moderate use of active learning]) spent 40% of class time on activities. Instructor 4 (identified hereafter as "Instructor 4-L" ["L" indicating a lecture-based delivery method]) spent less than 5% of class time on active learning (Table 1). The remaining class time (14–20%, 60%, and 95%, respectively, for Instructors 1 and 2, 3, and 4) was devoted to traditional, mostly passive, lectures. All four classes spent between three to six classroom hours on the topics of ecosystems and climate change. More instructional time was allocated to these subjects in classes devoted to active learning, because activities inherently take more time than lecturing. This difference in time on task might be considered a limitation of this study; however, other studies have shown that passive approaches to instruction (e.g., lecture only) are often ineffective for promoting learning, regardless of how much time is spent on a topic (Hake 1998; Knight and Wood, 2005; Minner *et al.*, 2010).

All four instructors attended both the 2008 and 2009 DQC PD program summer workshops and accepted the goal of improving students' understanding of carbon transformations by promoting principle-based reasoning; yet, due to time constraints, Instructor 4-L was not able to transition her course from a lecture-based to an active-learning format. Thus, only Instructors 1-AL, 2-AL, and 3-MAL shared the common objective of incorporating specific activities that targeted weaknesses in students' reasoning about carbon transformations based on pre-DQC results. More specifically, their intent was for students to apply, via activities, the principles of conservation of matter and energy when reasoning about the processes of photosynthesis, respiration, and biosynthesis. The various active-learning activities included clicker questions, think-pair-share exercises, conceptual questions, data-rich problem sets, and case studies. Brief summaries of the tar-

geted active-learning activities and their associated DQC assessment questions are provided in Supplemental Material 1, while Supplemental Material 2 provides an example of one of the in-class activities. Students were not assigned homework during the units related to carbon transformations.

Data Collection

At the beginning of the Fall 2009 semester, all four instructors administered two preinstruction DQCs to their students—"Keeling Curve" and "Grandma Johnson" (both available at www.biodqdc.org). Each DQC set had six to seven multiple-choice, true/false, or short-answer questions designed to assess student understanding of biological processes, specifically the flow of carbon and energy through ecosystems. Most multiple-choice and true/false questions also required a short-answer response explaining the answer choice, from which reasoning type (i.e., informal, mixed, or principled scientific reasoning) could be determined. These two sets of DQCs were selected because "Grandma Johnson" focuses on carbon cycling at the molecular and organismal levels, while "Keeling Curve" addresses carbon cycling at the organismal and ecosystem levels with links to climate change. Students were given as much time as needed to complete these questions within the constraints of the class period, but on average students took 15–25 min to complete each six-question DQC set. Taking the pre- and posttests was required for the course, but no points were given; instead, students were encouraged to try their best, as the data would be compared with data from other schools across the country. At the end of the semester, the same DQCs were administered in the same manner to assess student gains. Although students may have remembered some of the questions from the beginning of the semester, answers to the DQC assessments were not discussed in class.

Students' answers and written explanations to each pre- and postassessment question were scored by the instructor at the end of the semester. Student identity was removed before

coding and replaced with a four-digit label so that pre- and posttest results could be paired. Students who did not take a pre- or postassessment (e.g., absent that day) were excluded from the study. The three authors coded all of the responses, including the comparison class, using a validated coding scheme provided by the DQC developers. Instructors received training in administering and coding the DQCs during the DQC PD program workshops at the Ecological Society of America annual meetings in 2008 and 2009 (D'Avanzo and Anderson, 2008, 2009). Interrater reliability was determined by authors coding the same subset of questions independently; any discrepancies were discussed and negotiated to ensure at least 90% reliability. Furthermore, all of our data were sent to the DQC program organizers at Michigan State University and scores were validated independently for every 10th response and rescored when agreement was less than 90% (which was rare). The scoring rubric was based on four categorical ratings. A score of 4 indicated principle-based reasoning (students accurately apply principles of conservation of matter and energy), a score of 3 indicated mixed reasoning (students apply principles of conservation of matter and energy, although incompletely), a score of 2 indicated informal reasoning (no principle-based reasoning), and a score of 1 ("no data") indicated the student gave a nonsense answer, said "I don't know," skipped the question, or did not reach the question, or the response was illegible (Hartley *et al.*, 2011). For the complete coding rubric, see the Thinking Like a Biologist website (www.biodqc.org) Diagnosis Guides in Downloads.

Statistical Analyses

We analyzed results from seven of the 12 total questions on the DQC sets given to the students. The topics of these seven questions were directly related to the targeted active-learning exercises used in the classrooms. Of the five questions excluded from analysis, one was not related to the targeted active-learning exercises, and the other four had ambiguous wording identified only after the pretest was completed (the DQCs were in their final pilot stage at the time of this study).

Exploration of the data for alignment with normality showed that the pretest and posttest data sets did not display any significant skew or kurtosis (peakedness). Therefore, scores on the pre- and postassessments were treated as continuous data to allow for robust statistical tests. Because we were interested in identifying learning gains within and between groups, pre- and postassessment scores were compared on a by-question basis between instructors with analysis of variance followed by a Tukey least significant difference (LSD) post hoc comparison of the means. Analysis of variance was used to detect the effect, if any, of preassessment score, instructor, and the interaction of these variables on postassessment score, for each question. Paired *t* tests were used to detect significant differences in a group's mean pre- and postassessment scores for each question individually, and the set of questions (scores earned on the seven questions combined). Each of the seven questions in the set, regardless of question type (i.e., multiple choice, true/false, etc.) had a maximum code, or score, of four, thus making the set of questions worth a maximum of 28. Each student's score on the question set was converted to a percentage by dividing their sum of codes by the maximum sum of 28. Mean pre-

and posttest percentage were calculated for each group by obtaining the class average of these percentages.

Course average-normalized gains were calculated per question as the average of single-student normalized gains for each question (Hake, 2001). Normalized gain takes into account differences in student knowledge and measures the fraction of the available improvement that can be gained (Slater *et al.*, 2010). Course average-normalized gains by question were compared between groups with analysis of variance followed by a Tukey LSD post hoc comparison of the means using JMP version 7 software (www.jmp.com/software). Finally, course average-normalized gains were calculated by group for the set of questions (again, scores earned on the seven questions combined).

RESULTS

Preassessment

Prior to the intervention, mean pretest scores per question by class ranged from 1.7 to 3.2 (4.0 possible per question), with few significant differences ($p < 0.05$) per question between classes. Biology major students (Instructor 2-AL) scored significantly higher than non-biology major students (Instructor 1-AL) for only three out of seven questions on the preassessment. The comparison class (Instructor 4-L) initially scored significantly lower for three out of seven questions compared with the three other classes. Nevertheless, the data show that all of the students began their courses with informal to mixed reasoning skills.

Postassessment

Analysis of variance revealed that an instructor's teaching method had a highly significant effect ($p < 0.01$) on student reasoning following instruction for all questions. Students' initial reasoning skills, as indicated by pretest scores, was found to significantly influence ($p < 0.05$) their postinstruction reasoning for all topics, except biosynthesis in plants and carnivores. The interaction term (preassessment \times instructor) was not significant for any of the questions. Results of paired *t* tests showed that Instructors 1-AL and 2-AL, who employed active-learning methods at least 80% of the time, had significant increases ($p < 0.05$) in their students' posttest scores for six of the seven questions, and on the set of questions as a whole (Table 2). Instructor 3-MAL, who used active-learning

Table 2. Changes in student understanding of carbon-transforming processes following interventions^a

Instructor	Mean pretest (%)	Mean posttest (%)	p^b	DQC set <g>
1-AL	67	86	< 0.001	0.57
2-AL	75	91	< 0.001	0.63
3-MAL	68	79	0.023	0.25
4-L	52	55	0.122	0.05

^aAL: active learning; MAL: moderate active learning; and L: lecture.

^b*p* value from a paired-sample *t* test of pre- and posttest scores for the seven questions selected for analysis from the "Grandma Johnson" and "Keeling Curve" DQCs.

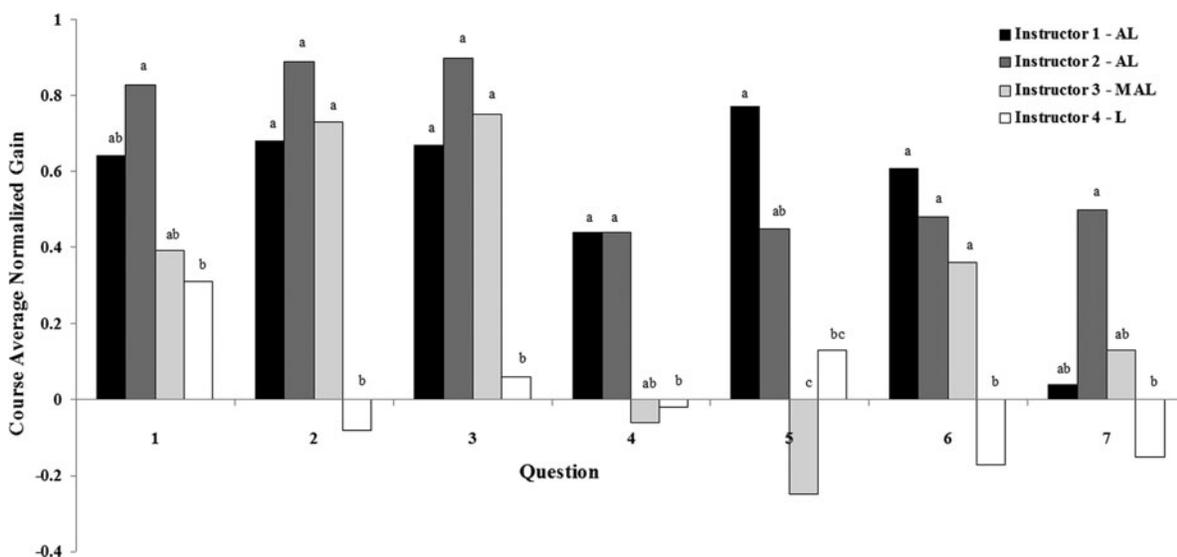


Figure 1. Changes in student reasoning about carbon transformations. Values graphed are course average normalized gains. Values for each question with different superscripts are significantly different at $p < 0.05$ between instructors. AL designates active learning, MAL moderate active learning, and L lecture interventions.

methods 40% of the time, had significant increases in her students' posttest scores for only two of the seven questions, yet her students' set of posttest scores increased significantly (Table 2). Posttest scores earned by Instructor 4-L's students, who received information in class from traditional lectures only, were not significantly different from pretest scores for any of the questions.

Course average-normalized gains and average posttest scores, representing the final reasoning levels attained by the students, were both significantly higher in courses incorporating active-learning strategies than in the class in which lecture was the dominant mode of instruction

(Figures 1 and 2 and Table 2). Individual contrasts between questions indicated that learning gains were significantly higher for students of the two AL instructors than for students of the non-active-learning instructor for all questions, except question 5 for Instructor AL-2's students and questions 1 and 7 for Instructor AL-1's students (Figure 1). Posttest scores allowed us to determine whether students in courses taught with targeted active-learning exercises reasoned more scientifically at the end of the course, which they did, for all questions (Figure 2). Although the pre/post-comparison within instructor groups showed significance ($p < 0.05$) for all but one question, average posttest scores from Instructor

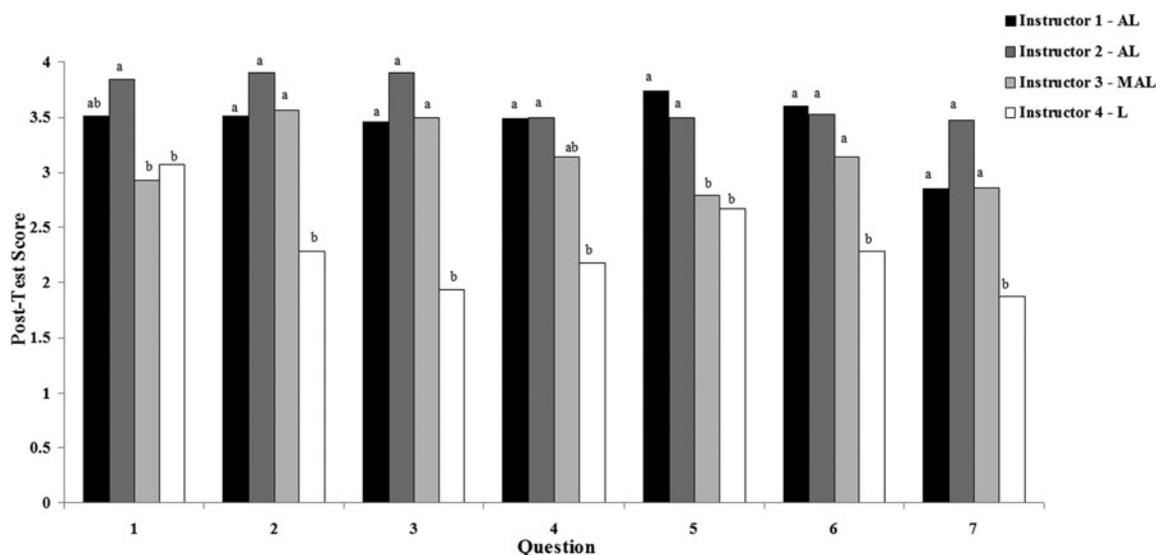


Figure 2. Postinstruction reasoning of students for topics related to carbon transformations. Values graphed are mean posttest scores. Values for each question with different superscripts are significantly different at $p < 0.05$ between instructors. AL designates active learning, MAL moderate active learning, and L lecture interventions. The highest score attainable was 4 and the lowest was 1.

1-AL's class were lower than Instructor 2-AL's class for four out of seven questions. This difference can be accounted for by the fact that Instructor 1-AL's students were non-biology majors and started at a lower reasoning level. For example, for questions 2 and 3 (cellular respiration and photosynthesis as a carbon sink), students averaged 2.3 and 2.4, respectively, on the preassessment for Instructor 1-AL, and 2.9 and 3.0, respectively, for Instructor 2-AL. Similarly, scores were still lower on the postassessment in Instructor 1-AL's class (average score of 3.5) when compared with Instructor 2-AL's class (average score of 3.9; Figure 2). Instructor 3-MAL used a mix of teaching strategies in the classroom and, for the most part, her students' posttest scores align with those of Instructors 1-AL and 2-AL for topics taught with targeted active-learning activities and with those of Instructor 4-L for topics conveyed by passive lecture (Figure 2).

DISCUSSION

The paired activities and assessment questions targeted specific ecological processes that play a prominent role in biology curricula and that are known to be difficult for students: generation (e.g., photosynthesis), transformation (e.g., biosynthesis), and oxidation (e.g., respiration). Principle-based reasoning by students would be indicated by explanations that trace carbon through multiple levels of organization and account for mass increases and decreases as a result of carbon transformations. Although the specific contexts for the assessment questions spanned primarily the ecosystem and global levels, accurate accounting for what occurs in these systems requires reasoning at the organismal, cellular, and subcellular scales as well.

Using specific in-class activities targeting known carbon transformation misunderstandings had a beneficial effect on student learning for both biology and non-biology majors in small and large classes (Figure 1); thus, our findings are not unique to one kind of classroom setting. Only marginal improvement, and not significantly so, was seen in the non-active-learning class after ecosystem ecology and the carbon cycle were covered in lectures (Figures 1 and 2). We also found that student performance on the preassessments was not significantly different between classes for the majority of the questions, which allowed the comparison of instructors utilizing different teaching methods to be more meaningful. Not surprisingly, however, performance on preassessments directly affected performance on postassessments for the majority of the questions. This is important to consider when interpreting student or class gains in knowledge, as subjects who reason at the mixed or scientific level have less to gain than those who began with informal reasoning approaches; therefore, normalized gains were calculated and compared. The significant effect of instructor on posttest score was desirable and indicative of the meaningfulness of our study design and results.

In the section below, we account for the results from each question that showed substantial gains, as well as those that did not. We chose to organize the discussion this way, because the goal of this project was to pair various activities with specific biological concepts that can be assessed with specific DQCs. Analyzing the data by category (generation, transformation, and oxidation) did not afford an opportu-

nity to describe the nuances in students' understandings and the influence of the specific activities on the knowledge and reasoning students applied.

Gains in Students' Reasoning by Topic

As suggested by the pretest scores, most of the students had difficulty correctly tracing carbon through the carbon cycle prior to instruction. Qualitative study of their responses revealed that they did not account for decomposers as a source of atmospheric carbon dioxide (questions 2 and 4) and the subsequent uptake by plants via photosynthesis (question 3; Supplemental Material 3). Instead, students reasoned that carbon dissolved into the soil and was directly taken up by plant roots. Although students' initial tracing of carbon at the ecosystem level was naïve, overall student improvement in the active-learning classrooms was high (Figure 1 and Table 2). We attribute this impressive increase to the in-class activity shown in Supplemental Material 2, because Instructors 1-AL, 2-AL, and 3-MAL implemented this activity and the classes of all three showed significant improvement compared with the comparison class (Figure 1 and Table 2). Instructor 2-AL's students had especially large gains for this topic, possibly because these biology majors had the opportunity to link together all of their disparate biology knowledge relevant to the carbon cycle learned from previous courses. Studies show that biology students, when asked, often can provide details about particular biological processes at one scale, but cannot make connections between processes at the cellular level and how they relate to organismal or ecosystem functioning (Maskiewicz, 2006). The activity used for this topic provided a scaffolded opportunity for students to reason about the possible paths carbon atoms could traverse longitudinally and to link together the various processes that would transform carbon along the way.

In-class activities were also effective when it came to explaining how plants acquire mass (question 5). Both Instructors 1-AL and 2-AL implemented activities that had students reason about results of classic plant experiments conducted by researchers such as von Helmholtz (Rabinowitch, 1971; Supplemental Material 1). Instructor 1-AL's students showed particularly impressive gains for this question (Figure 1) and, indeed, 77% of her students demonstrated principle-based reasoning on the posttest. Also noteworthy for this topic are the findings from Instructor 3-MAL's class, in which the students' level of understanding decreased, a result we attribute to a lab activity in which fertilizer was added to growing plants.² It is likely these students assumed that the increase in weight during the lab was the result of the fertilizer treatment, rather than photosynthesis, as illustrated by many students selecting the options "absorption of organic substances from the soil via the roots" or "incorporation of H₂O from the soil into molecules by green leaves" on this multiple-choice question. An improvement to this lab activity would be to require students to compare the dry mass of plants grown in a small, closed chamber with those grown in the open classroom to help students reason about how variation in available carbon dioxide produces differences in plant biomass.

²Although we are distinguishing laboratory activities from in-class active-learning activities, the results from Instructor 3-MAL's students warranted an explanation of how the laboratory activity may have influenced their results.

Students' understanding of mass change as a result of cellular respiration (question 4) was significantly higher for the classes of Instructors 1-AL and 2-AL compared with the comparison class (Figures 1 and 2). Several activities were specifically directed toward the objective of understanding that respiration releases carbon dioxide. In Instructor 2-AL's class, students also interpreted graphs showing the change in weight of fruit, mold, and fruit plus mold, and developed explanations for the different results. Even with this targeted activity, however, only 40% of Instructor 2-AL's students demonstrated principled reasoning. For those students who scored a "3" (mixed reasoning), most reasoned that mold growing on a piece of fruit would increase the overall weight (fruit plus mold), revealing that students were ignoring the fact that mold respire, releasing carbon dioxide and decreasing the overall mass. While Instructor 3-MAL used the "Cow Graphs" activity (Supplemental Material 2) in the classroom, an exercise that emphasizes carbon cycling at the ecosystem scale, she did not include an activity that specifically linked cellular respiration to mass change, an organismal-scale phenomenon, perhaps explaining her students' negative gain for this question (Figure 1).

Students' understanding that carbon is released by cellular respiration in plants (question 6) improved in the three classes implementing active-learning activities to address this topic. Instructors 1-AL and 2-AL showed substantial gains in the number of students demonstrating principle-based reasoning, whereas Instructor 3-MAL's students remained stable at a mixed reasoning level (Figures 1 and 2, question 6). The success of the students in the classes of Instructors 1-AL and 2-AL can be attributed to the von Helmholtz-related activities described previously and in Supplemental Material 1.

Limited Change in Students' Reasoning

The least overall gain in reasoning was seen in students' ability to trace carbon backward (e.g., question 7: "How much carbon in a coyote originally came from a plant?"). Although Instructors 1-AL, 2-AL, and 3-MAL implemented the "Cow Graphs" activity (Supplemental Material 2), only Instructor 2-AL implemented an additional activity asking students to trace where the carbon in a whale came from, and only her students showed improvement from pre- to posttest for this topic (Figure 1). While mean posttest scores for the students of Instructors 1-AL and 3-MAL were higher than those of Instructor 4-L (Figure 2), their scores did not increase significantly pre- to postinstruction for this question. These results suggest two things. First, tracing carbon backward is a more difficult task for students to reason through than tracing carbon forward from a decomposing organism (i.e., "Where was the carbon?" rather than "Where will the carbon go next?"). Second, instructors should implement targeted active-learning activities that provide opportunities for students to trace carbon in both directions, from carbon in the atmosphere to animal biomass and vice versa.

Differences between active- and passive-learning classes were less significant when addressing carbon cycling at the global level. In question 1, students were asked to link the rise and fall of atmospheric carbon dioxide levels with seasonal changes in photosynthesis and respiration. In active-learning classes, students were provided opportunities to reason about

the Keeling curve and respond to questions; this was followed by discussion. Although the classes of Instructors 1-AL and 2-AL improved significantly pre- to postinstruction, only one of these classes, that of Instructor 2-AL, was significantly different from the classes of Instructors 3-MAL and 4-L, in which students were shown the Keeling curve on a lecture slide and simply told the reason for the seasonal rise and fall of carbon dioxide (Figure 1).

CONCLUSION AND IMPLICATIONS

While this study provides insight into the effectiveness of active-learning techniques, we believe it is time to move past generic promotions of active-learning pedagogies and move toward the next generation of pedagogical reform: pairing targeted activities with diagnostic assessment strategies. Our results clearly reveal that targeted active-learning activities move student reasoning away from informal and mixed reasoning and toward scientific, principle-based reasoning. The two classes that received the most activities (those of Instructors 1-AL and 2-AL) had the most gain, while the 3-MAL class, with approximately 40% active learning, showed gain only for topics addressed with targeted activities. We do acknowledge, however, that the amount of time spent in class engaging with the concepts is also a factor in our results; instructors using active-learning approaches *did* spend more time on matter and energy transformations than the instructor in the comparison class. Nevertheless, we see this time allocation as a worthwhile and necessary investment, given that the gains from the lecture-only class were minimal at best.

Our data support other studies that show that passive instructional approaches, although efficient at covering a lot of material, are not effective at promoting meaningful learning (Hake, 1998; Beichner and Saul, 2003; Minner *et al.*, 2010). Several of these studies show the positive effects of active learning while controlling for class type and instructor (e.g., Knight and Wood, 2005). We found that, regardless of the student population, the two instructors that used 80% or more targeted active learning had greater learning gains. The differences in student populations are viewed as a strength of this study, and we account for any preinstruction differences by calculating normalized gain. Instructors 1-AL and 2AL used 80% or more active-learning strategies, and their classes showed great improvements in learning; yet group 1-AL was non-science majors, while 2-AL was upper-division biology majors. In contrast, the 3-MAL group, which experienced only 40% active learning, showed less gain than the other two groups, but more gains than the comparison group that experienced no active learning. Thus, our goal is not to claim equivalence among groups, but rather to show that active learning paired with diagnostic assessment of students' knowledge promoted learning in varied classrooms and with students at different levels.

While biology majors should be expected to have a firm, scientific grasp of the processes and transformations involved in carbon cycling, non-science majors should also be able to reason scientifically in order to participate as responsible citizens of a global community. Studies such as this one reveal that college students often have a limited understanding of terrestrial carbon dynamics, including the idea that plants absorb

carbon from decaying biomass in the soil rather than from the atmosphere. If students retain these conceptions, they cannot fully understand carbon-cycling phenomena, such as global climate change or the challenges in overcoming its effects. Therefore, faculty should consider addressing the connections between large-scale global phenomena and small-scale biological understanding by using conceptual assessments to diagnose students' current reasoning and inform course design (AAAS, 2010).

Faculty adoption of active-learning pedagogies in universities across the United States is varied, from those who deliver passive lectures only, through many who lecture and scatter a few in-class activities throughout their curriculum, to those who very rarely lecture and, instead, employ numerous activities throughout the semester that encourage students to do the intellectual work of reasoning about phenomena. Although some faculty may consider active learning impractical, because it reduces the amount of content that can be covered in a semester, the results from this paper suggest that pairing targeted active-learning exercises with related assessment tools is both effective and necessary. We encourage faculty to consider the evidence from our paper and think about using the results from diagnostic questions to create more effective classroom activities that target the reasoning level of their own populations of students. Funding to help educate faculty on using targeted active-learning activities and assessment tools (e.g., DQCs) should be given a high priority.

ACKNOWLEDGMENTS

This research was supported by NSF-DUE grants 0919992 and 0920186 to principal investigators Charlene D'Avanzo (Hampshire College), Charles Anderson (Michigan State University), and Alan Griffith (Mary Washington University). We appreciate the help given by Brook Wilke (Michigan State University) and Laurel Hartley (University of Colorado, Denver) in coding the DQCs. We thank the professors at Hampton University and James Madison University for sharing their DQC data with us. We also thank our students for their participation in this project and our institutions for their support. Finally, we thank Laurel Hartley, Charlene D'Avanzo, and the reviewers and editor for their substantive comments for improving this paper.

REFERENCES

- American Association for the Advancement of Science (AAAS) (2009). *Vision and Change in Undergraduate Biology Education: A View for the 21st Century*. www.visionandchange.org (accessed 28 January 2011).
- AAAS (2010). *Vision and Change: A Call to Action. A Summary of Recommendations*. <http://visionandchange.org/files/2011/03/VC-Brochure-V6-3.pdf> (accessed 28 January 2011).
- Anderson C, Sheldon TH, Dubay J (1990). The effects of instruction on college non-majors' conceptions of respiration and photosynthesis. *J Res Sci Teach* 27, 761–776.
- Armbruster P, Patel M, Johnson E, Weiss M (2009). Active learning and student-centered pedagogy improve student attitudes and performance in introductory biology. *CBE Life Sci Educ* 8, 203–213.
- Beichner RJ, Saul JM (2003). Introduction to the SCALE-UP (Student-Centered Active Learning Environment for Undergraduate Programs) Project. www.ncsu.edu/per/scaleup.html (accessed 1 February 2011).
- Brown JS, Collins A, Duguid P (1989). Situated cognition and the culture of learning. *Educ Researcher* 18, 32–41.
- Carlsson B (2002a). Ecological understanding 1: ways of experiencing photosynthesis. *Int J Sci Educ* 24, 681–699.
- Carlsson B (2002b). Ecological understanding 2: transformation—a key to ecological understanding. *Int J Sci Educ* 24, 701–715.
- D'Avanzo C (2008). Biology concept inventories: overview, status, and next steps. *Bioscience* 58, 1079–1085.
- D'Avanzo C, Anderson CW (2008). Using diagnostic question clusters (DQCs) to improve introductory ecology and biology teaching. Workshop for invited participants held at the Ecological Society of America 93rd Annual Meeting, Milwaukee, WI, 3–8 August 2008. www.esa.org/milwaukee/program.php (accessed 1 February 2011).
- D'Avanzo C, Anderson CW (2009). Using diagnostic question clusters (DQCs) to improve introductory ecology and biology teaching. Workshop for invited participants held at the Ecological Society of America 93rd Annual Meeting, Milwaukee, WI. Description retrieved from: <http://www.esa.org/albuquerque/program.php> (accessed 1 February 2011).
- Driver R, Squires A, Rushworth P, Wood-Robinson V (1994). *Making Sense of Secondary Science: Research into Children's Ideas*. London: Routledge.
- Ebert-May D, Batzli J, Lim H (2003). Disciplinary research strategies for assessment of learning. *Bioscience* 53, 1221–1228.
- Hake RR (1998). Interactive-engagement vs. traditional methods: a six-thousand-student survey of mechanics test data for introductory physics courses. *Am J Phys* 66, 64–74.
- Hake RR (2001). Suggestions for Administering and Reporting Pre/Post Diagnostic Tests. www.physics.indiana.edu/~hake/TestingSuggestions051801.pdf (accessed 1 February 2011).
- Hartley L, Wilke B, Schramm J, D'Avanzo C, Anderson A (2011). College students' understanding of the carbon cycle: contrasting principle-based and informal reasoning. *Bioscience* 61, 65–75.
- Hestenes D (1998). Who needs physics education research!? *Am J Phys* 66, 465–467.
- Johnson D, Johnson R, Smith K (1998). *Active Learning: Cooperation in the College Classroom*, 2nd ed., Edina, MN: Interaction Book.
- Knight JK, Wood WB (2005). Teaching more by lecturing less. *Cell Biol Educ* 4, 298–310.
- Maskiewicz A (2006). Applying DNR-based instruction to the learning and teaching of ecology. Paper presented at the Annual Conference of the National Association of Research in Science Teaching—San Francisco, California, April 2006.
- Michael J (2006). Where's the evidence that active learning works? *Adv Physiol Educ* 30, 159–167.
- Minner D, Levy A, Century J (2010). Inquiry-based science instruction—what is it and does it matter? Results from a research synthesis years 1984 to 2002. *J Res Sci Teach* 47, 474–496.
- Mohan L, Chen J, Anderson C (2009). Developing a multi-year learning progression for carbon cycling in socio-ecological systems. *J Res Sci Teach* 46, 675–698.
- National Research Council (NRC) (2000). *How People Learn: Brain, Mind, Experience, and School*. Washington, DC: National Academies Press.
- NRC (2003). *BIO2010: Transforming Undergraduate Education for Future Research Scientists*. Washington, DC: National Academies Press.
- NRC (2006). *Systems for State Science Assessment*. Committee on Test Design for K-12 Science Achievement, ed. M.R. Wilson and M.W. Bertenthal, Washington, DC: The National Academies Press.
- Nehm R, Reilly L (2007). Biology majors' knowledge and misconceptions of natural selection. *Bioscience* 57, 263–272.

Piaget J (1963/2003). Development and learning. *J Res Sci Teach* 40 (suppl), S8–S18.

Rabinowitch E (1971). An unfolding discovery. *Proc Natl Acad Sci USA* 66, 2875–2876.

Slater S, Slater T, Bailey J (2010). *Discipline-Based Education Research*, New York: W.H. Freeman.

Smith AC, Stewart R, Shields P, Hayes-Klosteridis J, Robinson P, Yuan R (2005). Introductory biology courses: a framework to support active learning in large enrollment introductory science courses. *Cell Biol Educ* 4, 143–156.

Songer CJ, Mintzes JJ (1994). Understanding cellular respiration: an analysis of conceptual change in college biology. *J Res Sci Teach* 31, 621–637.

Treagust DF, Chittleborough G, Mamiala T (2002). Students' understanding of the role of scientific models in learning science. *Int J Sci Educ* 24, 357–368.

Wilson CD, Anderson CW, Heidemann M, Merrill JE, Merritt BW, Richmond G, Sibley DF, Parker JM (2006). Assessing students' ability to trace matter in dynamic systems in cell biology. *Cell Biol Educ* 5, 323–331.