Article

Is Peer Interaction Necessary for Optimal Active Learning?

Debra L. Linton,* Jan Keith Farmer,[†] and Ernie Peterson*

*Department of Biology, Central Michigan University, Mount Pleasant, MI 48858; [†]Department of Biology, Schoolcraft College, Livonia, MI 48152

Meta-analyses of active-learning research consistently show that active-learning techniques result in greater student performance than traditional lecture-based courses. However, some individual studies show no effect of active-learning interventions. This may be due to inexperienced implementation of active learning. To minimize the effect of inexperience, we should try to provide more explicit implementation recommendations based on research into the key components of effective active learning. We investigated the optimal implementation of active-learning exercises within a "lecture" course. Two sections of nonmajors biology were taught by the same instructor, in the same semester, using the same instructional materials and assessments. Students in one section completed in-class active-learning exercises in cooperative groups, while students in the other section completed the same activities individually. Performance on low-level, multiple-choice assessments was not significantly different between sections. However, students who worked in cooperative groups on the in-class activities significantly outperformed students who completed the activities individually on the higher-level, extended-response questions. Our results provide additional evidence that group processing of activities should be the recommended mode of implementation for in-class active-learning exercises.

INTRODUCTION

Cooperative group-based active learning has become one of the benchmark techniques of the reformed, student-centered, introductory biology classroom. In these types of courses, students work together in groups during class to answer questions and solve problems related to key biological concepts. Many studies have shown the effectiveness of this technique in improving student learning (e.g., Udovic *et al.*, 2002; Knight and Wood, 2005; Armstrong *et al.*, 2007; Carmichael, 2009). Often, these types of studies of active learning compare student performance in cooperative group courses with performance in traditional lecture courses that include no

student-centered activities. These types of studies have been crucial to the widespread acceptance of the value of active-learning techniques. However, this type of design does not allow us to identify the key components of an effective active-learning exercise. This is an important line of investigation that requires our attention. For active learning to become more widely practiced, as recommended by Vision and Change (American Association for the Advancement of Science, 2011), we need to make it more accessible to all instructors. Andrews *et al.* (2011) found that active learning, as practiced by randomly selected "typical" college biology instructors, was not associated with greater student learning compared with traditional lecture. They point out that most studies of active learning are predominantly conducted in the classrooms of science education researchers who have had extensive experience and training in these techniques. Andrews and colleagues suggest that a "rich and nuanced understanding of teaching and learning" must be developed by "typical" instructors in order for them to be similarly successful with implementing active learning. We can help provide this understanding by identifying the key elements in effective active-learning implementation.

An active-learning exercise typically begins with the instructor presenting a task to the class. Students then work

on the task. After the students have finished their work, some sort of full-class processing typically takes place. Implementation of each of these stages could vary in important ways. Many of these variations are highlighted in the published meta-analyses of active learning in science (e.g., Prince, 2004; Michael, 2006; Ruiz-Primo *et al.*, 2011), and we have identified others based on our own experiences and conversations with colleagues. The nature of the question or problem itself varies widely in complexity. Students might be asked to answer a multiple-choice question (which themselves can vary in complexity and cognitive level), analyze graphical data, draw models, work through case studies, or perform a variety of other tasks. Students could work alone on the problem or in formal or informal cooperative groups. After group discussion, one team member might write the group's answer or each team member might write his or her own answer or students may not be asked to write at all. Full-class processing may involve the use of clickers. The instructor may ask the students to report and share their ideas with the class for discussion or may simply explain the answer to the students (Turpen and Finkelstein, 2009). Given all the possible differences in implementation, it is no surprise the outcomes are not consistent. Even those who are experienced and successful with active learning do not all do it the same way. However, if we can identify which components are crucial to the successful implementation of active learning, we can make it less of a risk for new instructors (or "old" instructors who want to try something new) to give it a try.

We are interested specifically in the effect of peer interaction on student learning and self-efficacy. Social interdependence theory (Johnson and Johnson, 2009) leads us to predict that peer interaction and cooperation will increase student learning when positive interdependence exists. In addition to this enhancement of learning, Johnson and Johnson (2009) have determined that cooperative learning is positively related to a variety of psychological attributes, such as confidence, independence, and self-esteem. Self-efficacy (students' belief in their ability to succeed) has been identified as an effective predictor of learning (Zimmerman, 2000). Therefore, we have included self-efficacy as a variable in this study.

Research in physics and chemistry education has found strong evidence that peer interaction can enhance learning. A meta-analysis (Gosser, 2011) of peer-led team learning (PLTL) in chemistry demonstrated increases in student learning in three different implementation formats. PLTL has also been shown to increase retention while maintaining rigor (Lewis, 2011). Peer instruction in physics has shown similar success (Crouch and Mazur, 2001) and is a widely accepted and practiced pedagogy in physics courses. However, Turpen and Finkelstein (2009) determined that differences in implementation of peer instruction can result in "disparate opportunities for students to practice conceptual reasoning, skills at talking physics, agency, and scientific inquiry" (p. 1).

Ruiz-Primo *et al.* (2011) identified differences in implementation of active learning and compared effect sizes in a meta-analysis of 197 comparative studies of innovations in science classrooms. In this analysis, they separated out different categories of innovations and their methods of implementation. The active-learning approach we have identified above would be categorized in their study as "conceptually oriented tasks + collaborative learning." They found this combination

to be the most common innovation infused into undergraduate science courses. They identified 41 published research studies implementing this approach (and collecting sufficient data for analysis) and found an effect size of 0.54. However, they also identified studies that implemented conceptually oriented tasks outside the structure of collaborative learning and found an effect size of 0.68. These results suggest greater gains when conceptually oriented tasks are implemented without cooperative learning.

We have to be careful when evaluating the effectiveness of peer interaction to make sure our results are not confounded by the effect of the activity that students experience. In the Ruiz-Primo *et al.* (2011) meta-analysis, studies that evaluated the use of conceptually oriented tasks versus traditional lecture are being compared with other studies that looked at cooperative activities versus traditional lecture. This is in contrast to a study design that could compare the use of *the same* conceptually oriented tasks implemented with and without collaboration. Studies of this type are needed for a variety of active-learning implementation issues.

Science education research is beginning to make progress on this agenda. Smith *et al.* (2009) found that peer discussion improved student performance on in-class clicker question sets. In 2011, Smith's research team expanded this research to include differential implementation of the clicker activities. They found that peer discussion followed by instructor explanation lead to greater learning gains than either peer discussion or instructor explanation alone. If this finding is confirmed, then we can add this to our shared definition of effective active learning. Discrete investigations of this type, which isolate the effect of specific components, are needed to provide the evidence that builds a shared definition of effective active learning.

We were particularly interested in the idea that the implementation of conceptually oriented tasks might be more (or at least as) effective without the incorporation of peer interaction. It is possible that the learning gains that occur because of an activity may be due primarily to the "time on task," individual thinking, or instructor explanation that occurs when practicing the application of recently learned biological concepts and may not require the cooperative group format. We do not believe that prior research has sufficiently disaggregated the effect of students' interaction with the *task* from the effect of students' interaction with their *peers*. We have attempted to do so here.

METHODS

Research Design

This research was approved by the Institutional Review Board (project #369307-1) at the university where it was conducted. Two sections of an introductory biology course for nonmajors were taught by the same instructor in the same semester, with back-to-back class meeting times. Both sections began the semester with 144 students enrolled; however, only 133 students in each section signed consent forms to participate in the research. The course included a 2-h laboratory experience each week. There were six laboratory sections of 24 students for each lecture section. The laboratory experiences, including all introductory materials and experimental

procedures, were the same for all sections. Consistency between sections was maintained by weekly meetings to train the graduate teaching assistants in facilitating each lab experience.

In both sections, 10- to 12-min lecture portions were alternated with in-class activities. In one section, students worked in cooperative groups on these activities. These groups were randomly assigned by the instructor, and each consisted of four students. The activities varied in complexity and included such tasks as: 1) drawing a model of a concept (e.g., relationship of nucleotide, gene, DNA, chromosome) or process (e.g., osmosis in freshwater vs. saltwater fish), 2) designing an experiment to test a hypothesis, 3) analyzing graphical data in context (e.g., changing phenotype frequency via natural selection), and 4) making a prediction or explaining a concept using evidence. In the other section, students completed the same activities but worked individually. All other aspects of course design, delivery, and assessment were held constant between the two sections. Full-class processing of each activity, facilitated by the instructor, occurred after the group or individual portion was completed. Varying balances of student and instructor explanation occurred during these discussions. However, in each case, the instructor ensured that the activity had been fully explained before moving on. The only difference between the two sections was the group versus individual processing of the activities.

For promoting positive group interdependence in the cooperative groups, one individual from each team was randomly selected to complete a five-question quiz at the end of each class period. All team members present in class that day were assigned the quiz grade this student earned. In the individual section, each student completed the quiz independently at the end of each class. The quiz score and points earned on the activity were compiled for an in-class score for each day.

At the beginning of the semester, students completed a demographic survey (e.g., gender, ethnicity, high school type, high school grade point average [GPA]), self-efficacy assessment instrument (Baldwin *et al.*, 1999), and key concepts pretest to evaluate the equivalency of the two groups. The key concepts pretest included one question selected from the textbook (Campbell *et al.*, 2006) test bank that the researchers and instructor agreed best represented the key concept of each of the 20 chapters covered during the semester, and modeled the format, quality, and cognitive level of questions that would be asked on the multiple-choice portion of the exams. The key concepts test and self-efficacy instrument were readministered at the end of the semester during the last regular class meeting. Student learning was assessed based on four exams, which consisted of both multiple-choice and extended-response questions. The texts of the essay questions that were analyzed for this research are included in Table 1.

Table 1. Exam essay questions and Bloom ratings

Exam 1: Matter, Bloom level 3	Dry ice is the solid form of carbon dioxide. A completely sealed airtight container contains a piece of dry ice with a mass of 10 g. As the temperature rises, the solid dry ice is converted directly into gaseous CO_2. After the dry ice is completely converted to a gas, how has the amount of mass (matter) inside the container changed? Explain your answer using the vocabulary related to the nature of matter.
Exam 1: Protein production, Bloom level 2	Imagine a cell producing a protein that will be exported from the cell. Trace the pathway of production from the gene in the nucleus, to production of the protein, through the endomembrane system to the protein's export from the cell. Name the organelles involved and what happens at each step.
Exam 2: Photosynthesis and cellular respiration, Bloom level 3	Explain how a carbon atom in a molecule of glucose in one of your muscles could become part of a glucose molecule in a tomato plant in your garden. As you trace the pathway, name both the processes involved and how the carbon atom exits and enters any organisms involved.
Exam 3: Natural selection, Bloom level 4	On the Hawaiian island of Kauai, there is a species of cricket named *Teleogryllus oceanicus*. In the 1980s male crickets in this population could be heard chirping at night, attracting the females for mating. Almost all of the males chirped, but there was a small proportion of males that could not chirp because their wings lacked the structures needed for chirping. In the 1990s a fly species (*Ormia ochrace*) settled on the island. The flies implant their larvae in the bodies of crickets. The larvae use the crickets as food, and the crickets eventually die. The chirping males make so much noise that they are easy for the flies to locate, so they suffered the most mortality. By 2003, the proportion of nonchirping males had risen to more than 90%.
	a. How did the allele for nonchirping first come into existence?
	b. Explain how the process of natural selection resulted in the increase in the proportion of non-chirping males.
	c. If the flies were removed in 2004, predict what would happen to the proportion of chirping males in the cricket population over the next several generations. Explain your prediction.
Exam 4: Carbon cycling and energy flow in ecosystems, Bloom level 3	This drawing [food chain of plant → grasshopper → mouse → snake → hawk] represents a food chain. Organic matter present in each organism represents both stored energy and stored carbon. The arrows represent movement of energy and carbon atoms through the food chain. Answer the following questions based on this drawing.
	a. Where did the energy stored in the organic matter originally come from and by what process did it enter the food chain?
	b. Explain the shortest route that a carbon atom in a glucose molecule in one of the mouse's cells could take to end up back in a glucose molecule produced by the flower. For each step, name the processes involved and the form that the carbon will be in.
	c. What is the final fate of the energy stored in the mouse's glucose molecule?

Data Analyses

We ran chi-square analyses of each item on the student profile survey, comparing students in the cooperative group section with students in the individual activity section. Only students for whom we had key concepts pretests and all four sets of exam data were used in this analysis, as these are the only students we would be comparing in later analyses. The sample sizes for this analysis and all exam analyses are 102 students in the cooperative section and 98 in the individual section.

For an overall look at the effects on self-efficacy, scores on the self-efficacy instrument were analyzed by converting Likert responses to numerical values (totally confident = 5, through not at all confident = 1) and summing values to give a composite score for each student. Only students who completed both the presurvey and postsurvey were included in the self-efficacy analyses. This resulted in sample sizes of 92 students in the cooperative section and 95 students in the individual section. Pretest scores were compared between sections using a *t* test. Posttest scores were analyzed by analysis of covariance (ANCOVA), with scores on the pretest as the covariate.

Students' percent scores on the key concepts pretest were compared between treatments using a *t* test. Posttest scores were compared using ANCOVA, with pretest scores as the covariate. Again, only students who completed both the pretest and posttest were included in the analyses. For this analysis, the sample sizes were 86 students in the cooperative section and 92 students in the individual section.

We rated each exam item by Bloom level, using the original Bloom's taxonomy (1956). Three members of the research team rated each item. Any disagreements were discussed to reach consensus. Average Bloom levels were calculated for the multiple-choice and essay sections of each exam.

Students' percent scores on only the multiple-choice portions of the exams were compared between sections using ANCOVA, with percent score on the key concepts pretest as the covariate.

Essay responses on the exams were assessed separately from the multiple-choice questions. Essay responses were coded for specific key correct concepts based on a coding rubric developed by the research team. The coding rubrics were developed by the lead investigator, then reviewed by the rest of the research team, and edited based on their feedback. The rubrics were then piloted on a random sample of student work and further modified. The final rubrics are shown within the data figures for the essay analyses below. Essay answers were deidentified before coding, so raters were not aware of a student's treatment status while coding that student's answers. Two raters coded each essay, and disagreements were settled by discussion. The total number of correct concepts included in student responses was compared between groups by ANCOVA, using the percent score on the key concepts pretest as the covariate. Individual concepts were analyzed for differences in frequency between sections using chi-square analyses.

RESULTS

Demographic Data

No significant demographic differences were detected between the two sections in gender, major, highest level of par-
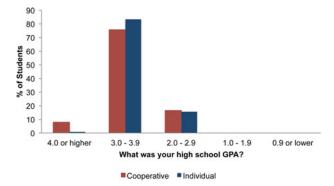

Figure 1. Self-reported high school GPA distributions of students in cooperative and individual sections. Distributions are significantly different ($p = 0.04$). $n = 102$ cooperative, 98 individual.

ents' education, high school type (urban/suburban/rural or public/private/homeschool), previous biology experience, or biology GPA. However, a significant difference was found for overall high school GPA ($p = 0.04$), with students in the cooperative group section self-reporting slightly higher overall high school GPA (Figure 1). On the basis of this information, we decided to use scores on the key concepts pretest as a covariate in analyses to minimize any inherent differences between sections. The GPA data were not used as a covariate, because they were categorical (students selected from GPA ranges) and also because they were less reliable self-reported data.

Self-Efficacy

Scores on the self-efficacy pretest were not significantly different ($p = 0.63$) between sections. Scores on the posttest were also not significantly different ($p = 0.35$; Figure 2). However, both the cooperative ($p = 0.00076$) and individual ($p = 0.016$) sections showed significant increases in self-efficacy from pretest to posttest (Figure 2).
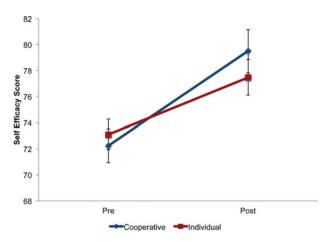

Figure 2. Average self-efficacy scores, pretest and posttest for both treatments. Bars represent SE. Means are significantly different pretest vs. posttest for both the cooperative ($p = 0.00076$) and individual ($p = 0.015$) sections, but not cooperative vs. individual on either the pretest ($p = 0.63$) or posttest ($p = 0.35$). $n = 92$ cooperative, 95 individual.
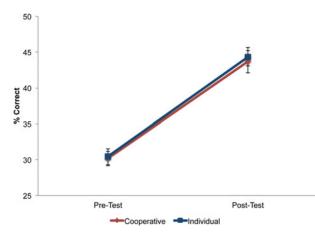
Figure 3. Average scores on the key concept test (pre and post) for both treatments. Bars represent SE. Means are significantly higher on the posttest compared with the pretest for both the cooperative ($p < 0.0001$) and individual ($p < 0.0001$) sections but are not significantly different between the sections on either the pretest ($p = 0.86$) or posttest ($p = 0.73$). $n = 86$ cooperative, 92 individual.

Key Concepts Test

Neither pretest scores ($p = 0.86$) nor posttest scores ($p = 0.73$) were significantly different between treatments. However, scores on the key concepts test were significantly higher on the posttest than the pretest for both the cooperative ($p < 0.0001$) and individual ($p < 0.0001$) sections (Figure 3).

Bloom's Levels

Average Bloom levels on the multiple-choice portions of the exams ranged from 1.40 to 1.61 (Table 2). The essay questions analyzed for this research ranged from Bloom's level 2 (comprehension) to level 4 (analysis), with exam averages ranging from 2.5 to 4.0 (Table 2). The ratings for individual essay questions are shown in Table 1.

Exams

There were no significant differences between treatments on the multiple-choice parts of any of the four exams (Figure 4). However, students in the cooperative group section included significantly more correct concepts in their answers to the essay questions on all four exams (Figure 5). The p values for these analyses are reported in the figure legends. An example of coded student work is presented in Table 3.

Table 2. Average Bloom levels for multiple-choice and essay questions on exams

Exam	Multiple-choice Bloom level	Essay Bloom level
1	1.49	2.5
2	1.44	3.0
3	1.61	4.0
4	1.40	3.0

Fifty-nine separate chi-square analyses were run on individual items from the exam essay rubrics. Seventeen items were found to have significant differences between the cooperative and individual sections. In each of those 17 cases, the cooperative section showed a significantly higher percentage of students including that item in their essay answer than did the individual section.

On exam 1, two essay questions were analyzed focusing on the major concepts from the chemistry and cell unit. Students in the cooperative section were more likely to reference the law of conservation of matter (Figure 6) in their explanation of the nature of matter and were also more likely to explain the function of the Golgi apparatus in the protein production (and modification) pathway (Figure 7).

On exam 2 (Figure 8), the key essay question related to the movement of carbon atoms during photosynthesis and cellular respiration. Five items were significantly different between sections. Students in the cooperative group section were significantly more likely to explicitly state that carbon dioxide is released from the body and subsequently taken in by a plant. They were also significantly more likely to identify the processes of cellular respiration and photosynthesis as the driving forces for this movement of carbon in ecosystems.

On exam 3 (Figure 9), the major essay featured natural selection. There were six concepts that students in the cooperative section included significantly more frequently than did students in the individual section. These concepts included genetic variation, differential survival, differential reproduction, and correct use of the term "fitness" in two different contexts.

The essay question analyzed for exam 4 focused on the flow of energy and cycling of carbon in ecosystems. Four concepts showed significant differences between sections (Figure 10), with students in the cooperative section being more likely to describe the production of carbon dioxide by cellular respiration and to explicitly state that the carbon for photosynthesis comes from the atmosphere and enters a plant through stomata and that the eventual fate of energy in ecosystems is to be lost as heat.

DISCUSSION

In this study, peer interaction was not shown to have a significant influence on students' overall self-efficacy. Self-efficacy increased from the beginning to the end of the semester, but students in the cooperative groups did not realize any greater overall increase than students who worked individually. This is in contrast to data from Prince's (2004) meta-analysis, which found several studies showing improved "self-esteem" from collaborative and cooperative learning compared with traditional lecture. Citing the work of Johnson *et al*. (1998) and a meta-analysis by Springer *et al*. (1999), they found increases in students' self-esteem and attitudes, with effect sizes of 0.29–0.67, in collaborative or cooperative groups. Similarly, Fencl and Scheel's 2005 study showed that collaboration can make positive contributions to self-efficacy. However, recall that our "noncooperative" treatment was not traditional lecture and self-efficacy did significantly increase in both sections. Individual student engagement with conceptually oriented tasks showed as much of an impact on self-efficacy as did
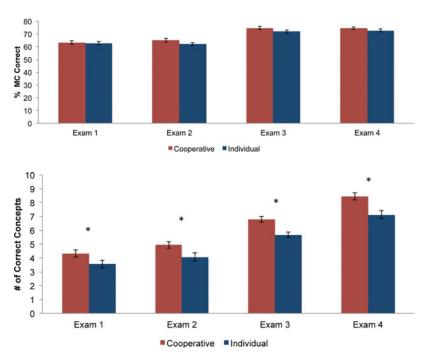
Figure 4. Class means (unadjusted) for percent correct on multiple-choice (MC) portions of exams. There were no significant differences between treatments based on ANCOVA. Exam 1 ($p = 0.89$); exam 2 ($p = 0.13$); exam 3 ($p = 0.14$); exam 4 ($p = 0.29$). $n = 102$ cooperative, 98 individual.

Figure 5. Class means (unadjusted) for number of correct concepts included in essay writings. Means were significantly different (∗) for all exams. Exam 1 ($p = 0.043$); exam 2 ($p = 0.023$); exam 3 ($p = 0.00018$); exam 4 ($p = 0.00078$). $n = 102$ cooperative, 98 individual.

group processing of those same activities. This suggests that peer interaction is not the primary driver of the increases in self-efficacy that have been seen in active learning.

Our results did not show any effect of peer interaction on students' ability to correctly answer multiple-choice questions, either on the key concepts posttest or on the exams. Previous studies have shown that peer discussion can improve student performance on multiple-choice questions. In the study conducted by Smith *et al.*, (2009), students discussed multiple-choice questions in cooperative groups and showed improvement on later multiple-choice questions on the same concepts. However, in their design, the improvement was not compared with any other mode of interacting with the questions. Smith and colleagues' research was investigating whether peer discussion can help students learn, while we are asking whether peer discussion is necessary to maximize learning during active-learning activities or whether students could interact in other ways and receive equal or greater benefit. Recall that in our study, full-class processing of the activities facilitated by the instructor always occurred after the group or individual work. This was not the case in Smith *et al.*'s 2009 paper. Later research by Smith and colleagues (2011) compared student performance on clicker questions under three scenarios: peer discussion alone, instructor explanation alone, and peer discussion followed by instructor explanation. This research is more analogous to our study. In

our cooperative sections, group work was always followed by instructor-facilitated explanation. In the individual section, there was only instructor-facilitated explanation. Smith *et al.* (2011) found that peer discussion followed by instructor explanation led to greater increases in performance than did instructor explanation alone. On the multiple-choice part of our research, this was not the case.

The comparison of student performance on essay questions tells a different story. Students in the cooperative section included more correct concepts in their answers than did students in the individual section. When individual concepts were compared, many significant differences were found, and in all cases in which a significant difference was found, the cooperative group section had a higher proportion of students including these key concepts in their essay answers. The fact that there were no differences seen in multiple-choice performance makes this trend stronger, as it (along with the use of key concepts pretest scores as covariates in all analyses) weakens any argument that students in the cooperative section were simply better students. The evidence is strong that working in cooperative groups and engaging in peer discussion resulted in greater performance by students on the higher-level assessments. The multiple-choice questions on the exam required primarily low-level cognitive skills as rated using Bloom's taxonomy. This is a common pattern in introductory biology courses (Momsen *et al.*, 2010). However,

Table 3. Examples of student work and coding for exam 4 essay question

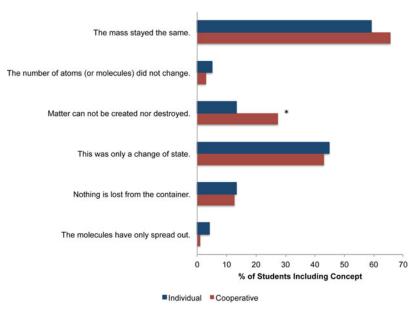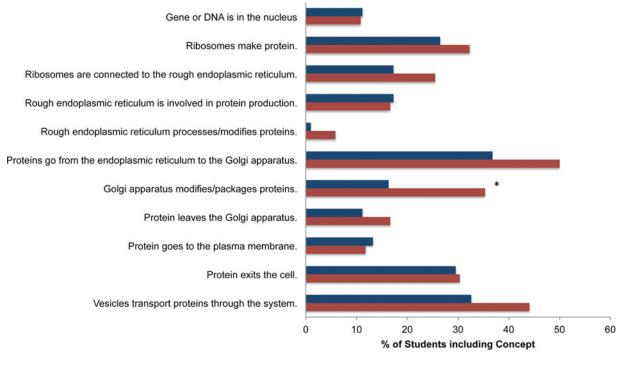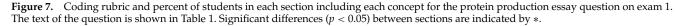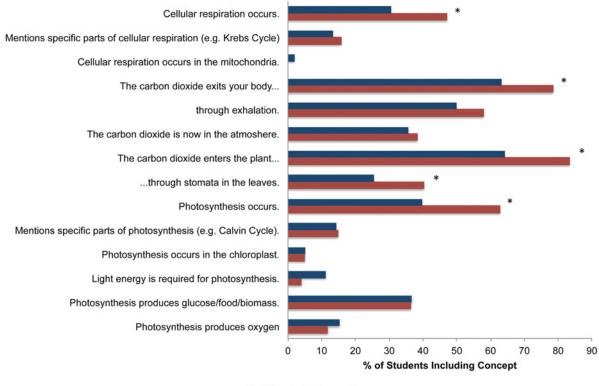| More complex answer: 7 correct concepts	The mouse will use the glucose molecule in cellular respiration. As the product of cellular respiration, the CO_2 is released into the atmosphere. The plant will then take it in from the atmosphere through the stomata in the leaves. Then the plant will use it in photosynthesis and the carbon atom will become part of a glucose molecule in the flower.
Less complex answer: 4 correct concepts	The mouse releases the carbon atom with cellular respiration. The plant then takes in the carbon with photosynthesis.

Figure 6. Coding rubric and percent of students in each section including each concept for the matter essay question on exam 1. The text of the question is shown in Table 1. Significant differences ($p < 0.05$) between sections are indicated by ∗.

the essay questions did make higher cognitive demands on students. The in-class activities used in the course varied in cognitive level but included many activities that would be rated at Bloom's levels 3–5. We suggest that the benefits that arise from peer interaction are most likely seen at these cognitive levels. Students do not need their peers to help them

memorize a fact and repeat it back on the exam. This finding aligns with Johnson and Johnson's (2009) predictions based on social interdependence theory. They identified conditions under which individual efforts might be most appropriate. One of the conditions they identified was "Unitary, nondivisible, simple tasks need to be completed, such as the learning

Figure 7. Coding rubric and percent of students in each section including each concept for the protein production essay question on exam 1. The text of the question is shown in Table 1. Significant differences ($p < 0.05$) between sections are indicated by ∗.

Figure 8. Coding rubric and percent of students in each section including each concept for the photosynthesis/respiration essay question on exam 2. The text of the question is shown in Table 1. Significant differences ($p < 0.05$) between sections are indicated by ∗.
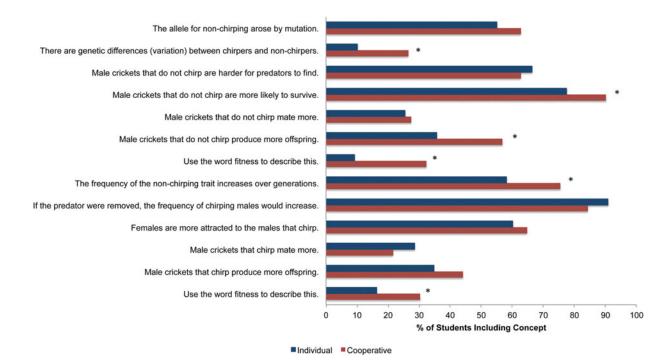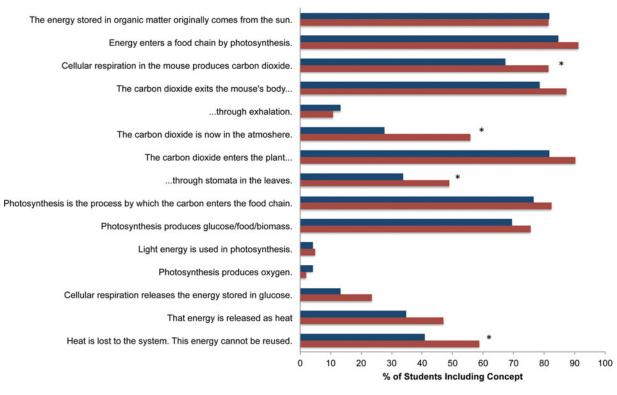

Figure 9. Coding rubric and percent of students in each section including each concept for the natural selection essay question on exam 3. The text of the question is shown in Table 1. Significant differences ($p < 0.05$) between sections are indicated by ∗.

Figure 10. Coding rubric and percent of students in each section including each concept for the carbon cycle essay question on exam 4. The text of the question is shown in Table 1. Significant differences ($p < 0.05$) between sections are indicated by ∗.

of specific facts or the acquisition or the performance of simple skills" (p. 370). Our results provide evidence in support of this hypothesis.

However, cognitive tasks that require application, analysis, and synthesis benefit from the experience students have discussing these concepts with their peers. On the basis of the example clicker questions included in Smith *et al.*'s 2011 paper, we do not see their data as contradictory to ours. Their multiple-choice questions were not the simple knowledge and understanding (Bloom's levels 1–2) questions typical of introductory biology exams, but required higher-level processing skills. From this perspective, the two studies support each other. Similarly, the ConcepTests and analogous exam questions used in physics peer instruction are not questions that can be answered by memorization (Crouch and Mazur, 2001). The proven effectiveness of peer instruction in these conditions further supports our findings. On higher-level tasks, group discussion followed by instructor-facilitated explanation leads to optimal student learning.

Another trend that has captured our interest relates to the types of concepts that show significant differences between sections on the essay rubrics. For many of the items on the rubrics, there were key terms that raters were instructed to look for, for example, "conservation of matter," "cellular respiration," "photosynthesis," "stomata," and "fitness." Other items focused on concepts that did not have specific terms associated with them. Many of the items that targeted key words showed significantly higher inclusion by students in

the cooperative group section. The current study was not designed to test this idea, so we cannot say anything definitive about this but suggest that it could be an area for further study. Cooperative group discussions may encourage students to use correct vocabulary, as it improves communication skills by encouraging scientific discussion between peers.

In conclusion, the evidence we have presented supports the assertion that "peer interaction is necessary for optimal active learning." However, the beneficial effects of peer interaction may only be seen on higher-order activities/assessments. The more complete answers written by students in the cooperative section suggest that it is not just the activity itself that improves learning but that peer interactions play a crucial role in promoting students' understanding or at least students' ability to explain their understanding. Therefore, as we begin to develop a shared definition of effective active learning, we propose that peer interaction, followed by instructor-facilitated explanation (during full-class processing), become a key component of that definition. While peer interaction followed by instructor-facilitated explanation is already a typical component of how most active learning is implemented, the framework of "scientific teaching" (Handelsman *et al.*, 2004) requires us to make these decisions based on evidence. This study, coupled with Smith *et al.* (2011), builds on the foundational research on cooperative learning to provide strong evidence in favor of this approach within a conceptually oriented task framework utilizing higher-level

activities and assessments. We renew our call for more research of this type to identify other key components of effective active learning, so we can more easily and broadly disseminate this approach in a manner that will allow all instructors to become effective with it.

REFERENCES

American Association for the Advancement of Science (2011). Vision and Change in Undergraduate Biology Education: A Call to Action, Washington, DC.

Andrews TM, Leonard MJ, Colgrove CA, Kalinowski ST (2011). Active learning not associated with student learning in a random sample of college biology courses. CBE Life Sci Educ *10*, 394–405.

Armstrong N, Chang S, Brickman M (2007). Cooperative learning in industrial-sized biology classes. CBE Life Sci Educ *11*, 17–25.

Baldwin JA, Ebert-May D, Burns DJ (1999). The development of a college biology self-efficacy instrument for nonmajors. Sci Educ *83*, 397–408.

Bloom BS (1956). Taxonomy of Educational Objectives: The Classification of Educational Goals. Handbook I: Cognitive Domain, New York: Longmans, Green.

Campbell NA, Reece JB, Simon EJ (2006). Essential Biology with Physiology, 2nd ed., New York: Benjamin Cummings.

Carmichael J (2009). Team-based learning enhances performance in introductory biology. J Coll Sci Teach *38*(4), 54–61.

Crouch CH, Mazur E (2001). Peer instruction: ten years of experience and results. Am J Phys *6*, 970–977.

Fencl H, Scheel K (2005). Engaging students: an examination of the effects of teaching strategies on self-efficacy and course climate in a nonmajors physics course. J Coll Sci Teach *35*(1), 20–24.

Gosser DK (2011). The PLTL boost: a critical review of research. Progressions *14*(1).

Handelsman J, Ebert-May D, Beichner R, Bruns P, Chang A, Gentile J, Lauffer S, Stewart J, Tilghman SM, Wood WB (2004). Scientific teaching. Science *304*, 521–522.

Johnson DW, Johnson RT (2009). An educational psychology success story: social interdependence theory and cooperative learning. Educ Res *38*, 365–379.

Johnson DW, Johnson RT, Smith KA (1998). Cooperative learning returns to college: what evidence is there that it works? Change *30*, 26–35.

Knight JK, Wood WB (2005). Learning more by lecturing less. Cell Biol Educ *4*, 298–310.

Lewis SE (2011). Retention and reform: an evaluation of peer-led team learning. J Chem Educ *88*, 703–707.

Michael J (2006). Where's the evidence that active learning works? Adv Physiol Educ *30*, 159–167.

Momsen JL, Long TM, Wyse S, Ebert-May D (2010). Just the facts? Introductory undergraduate biology courses focus on low-level cognitive skills. CBE Life Sci Educ *9*, 435–440.

Prince M (2004). Does active learning work? A review of the research. J Eng Educ *93*, 223–231.

Ruiz-Primo MA, Briggs D, Iverson H, Talbot R, Shepard LA (2011). Impact of undergraduate science course innovations on learning. Science *331*, 1269–1270.

Smith MK, Wood WB, Adams WK, Wieman C, Knight JK, Guild N, Su TT (2009). Why peer discussion improves student performance on in-class concept questions. Science *323*, 122–124.

Smith MK, Wood WB, Krauter K, Knight JK (2011). Combining peer discussion with instructor explanation increases student learning from in-class concept questions. CBE Life Sci Educ *10*, 55–63.

Springer L, Stanne M, Donovan S (1999). Effects of small-group learning on undergraduates in science, mathematics, engineering, and technology: a meta-analysis. Rev Educ Res *69*, 21–52.

Turpen C, Finkelstein ND (2009). Not all interactive engagement is the same: variations in physics professors' implementation of peer instruction. Phys Rev ST Phys Educ Res *5*, 020101.

Udovic D, Morris D, Dickman A, Postlethwait J, Wetherwax P (2002). Workshop biology: demonstrating the effectiveness of active learning in an introductory biology course. BioScience, *52*, 272–281.

Zimmerman BJ (2000). Self-efficacy: an essential motive to learn. Contemp Educ Psychol *25*, 82–91.