Group Problem-Solving Sessions Trump Lecture: Results of an Experiment in Teaching

Beverly Clendening, Ph.D
Department of Biology
Hofstra University
Hempstead, NY 11549

Abstract
Despite mounting evidence that lectures alone are not the most effective way to help students gain knowledge, faculty in undergraduate science programs still use more student-centered teaching formats infrequently. Numerous barriers to change exist, including the assumption that less science content can be covered when student-centered teaching formats are adopted. In this study I show that the replacement of one hour of lecture per week in an introductory cell biology and genetics course with a group problem-solving session resulted in no reduction in science content covered and a significant increase in correct responses to knowledge, comprehension and application/analysis assessment items on the final exam.

Introduction

Over the past 15 years several high profile task forces have examined undergraduate science teaching practices and made recommendations for improvement based on research about how people learn ((NSF 1996, NRC 1997, Boyer Commission 1998, NRC, 2000, NRC 2003). One of the principle recommendations has been to shift the emphasis in the science classroom from the passive learning that occurs in lectures to more active, student-centered forms of learning. A growing number of innovative approaches to undergraduate science education have been developed to promote active learning in the classroom. In most cases, research has shown these approaches to be effective in promoting student learning (Hake 1998, NRC 2000, Crouch and Mazur 2001, Born 2002, Knight and Wood 2005, Handelsman 2008; Carmichael 2009). Those that incorporate a small-group learning component have been particularly effective (Springer et al. 1999, Johnson et al. 2000; but see Dochy 2003). Despite the evidence of its effectiveness, faculty in undergraduate science programs still use active learning in the classroom infrequently for a variety of reasons (DeHaan 2005, Handelsman 2007).

Frequently cited barriers to change in undergraduate science teaching include time constraints, lack of support from administration and concern about the ability to cover sufficient content (DeHaan 2005). The concern that time devoted to active learning modules reduces the time available to teach students the facts they need to know as they enter jobs in industry and doctoral program in science or medicine is prevalent. While it is true that competence in any area of science requires a “deep foundation of factual knowledge” (NRC 2003), there is now considerable evidence to show that a variety of active learning strategies help students gain knowledge more effectively than lectures (Springer et al. 1999, NRC 2000, Knight and Wood 2005, Handelsman 2008, Carmicheal, 2009). The results of the study reported here add additional support to claims that comprehension and analytical skills increase when active learning (in this cases, in the form of group problem solving) is added to classroom activities.

Recently I have made a series of changes in the course format of a freshmen level introductory to cell biology and genetics course with the goal of improving the students’ ability to solve problems and to apply the concepts they learn to novel situations. Originally, the lecture portion of the course included only short active learning components such as think-pair-share, minute papers and concept mapping. Students taught using this format typically scored very poorly on final exam questions that required application and analysis as defined by Bloom (1956). In an attempt to improve the students' problem solving and analytical skills, I added homework problems sets that were done individually and graded. This strategy resulted in insignificant gains in the students' ability to answer application and analysis questions on the final exam. Because small group learning sessions have been particularly effective in improving higher-level reasoning in a number of different settings (Johnson et al., 2000), I next changed the course format by replacing one day of lecture per week with a group problem-solving session. The goal of this change in course format was specifically to improve the students' analytical skill and ability to apply knowledge to novel situations. I hypothesized that students taught using the lecture plus small group problem-solving format would be better able to answer application and analysis type final exam questions than students who were taught using the lecture plus individual problem-solving format. Because this course is one of the core introductory biology courses in my department I am expected to teach a specific set of concepts. Therefore, I could not change the breadth of content I presented but I could attempt to assess how much of what was presented was actually learned by the students. Therefore, I also tested the hypotheses that substitution of time spent lecturing with small group problem solving would result in no reduction in the students' content knowledge. I tested this hypothesis by comparing scores on final exam knowledge questions between students taught using the lecture and those taught with the lecture plus group problem solving formats.

The replacement of part my lecture time with group problem-solving sessions resulted in a significant improvement in final exam scores not only on application and analysis assessments but also on assessments of factual knowledge and comprehension. In addition, substituting group work for lecture time resulted in no reduction in the breadth of concepts covered in class. In fact, I was able to "cover" more information using the new format.

Method

The course and students

This study was conducted over a four-year period (2005-2008) in one section of an introductory cell biology and genetics course. For each year of the study, the section of the class used for the study was offered during the fall term and included 42-48 first semester freshmen. These students had the same range of innate ability (based on SAT/ACT scores) and had taken high school biology and chemistry and at least 3 years of high school mathematics. The number of students who had taken advanced biology in high school was nearly equivalent in each year of the study. None of the students in the study had earned college credit for biology taken in high school. The male: female ratio was different in each year, however, the ratio of male: female in the combined sections before (23 male: 50 female) and after (26 male: 56 female) the change in teaching format was nearly the same. The course consisted of three 50-minute lectures sessions, one 3-hour laboratory and one 1-hour recitation each week. The recitation was used to discuss laboratory results. For each year of the study, the lectures, laboratories and recitations were scheduled at the same time and the same instructor (the author) was assigned to all lecture, laboratory and recitation sections. The only changes in laboratory and recitation content over the four-year period of this study were minor revisions in laboratory procedures. The same textbook was used for all four years and the content covered was the same with the exception that under the lecture plus group problem-solving I was able to cover one additional topic (cell signaling) in much greater depth at the end of the semester.

Class format changes

In the first two years included in this study (2005 and 2006), for each learning unit, I assigned sets of key questions, problems and case studies as homework to help students learn the concepts being presented in lecture. The problems and case studies were also designed to improve students' problem-solving and analytical skills. Key questions were used as an organizing tool for the lectures. Students were not required to submit answers to the key questions for grading. The problems and cases were assigned 5-7 days in advance of the class in which they were discussed and students were required to hand in the assignment before class began on that day. During class we discussed the case studies and some of the more difficult problems. Problem sets were graded and this homework constituted 16% of the students' final grade.

In the second two years of this study (2007 and 2008), I retained the individual key questions as an organizing tool for my lectures. However, I reduced lecture time by approximately 30%. Once per week, in lieu of lecture, the students worked in groups to complete and review the problem sets and cases that had been assigned 5-7 days in advance. Students were required to submit an electronic copy of their individual answers to the problem set before class began. This homework was not graded, however, students earned credit for completed work that was handed in before class. This homework constituted 6% of the students' final grade. The assignment of credit for homework provided an incentive for the students to work on the problems before the in-class group sessions. In addition, students who failed to submit individual answers to the problems prior to the group sessions received a maximum of one-half of the points awarded their group for the group report.

The in-class group work was organized following the Process Workshop (POGEL) model (David Hanson, SUNY, Stony Brook). Students worked in groups of four to discuss and finalize answers to the problems and cases. Each student was assigned a number on the first day of class and groups for each session were formed using a random number generator. Therefore, during each session, each student worked with a different group of students. During each group session, each student selected a task within his/her group. Tasks included leader, reporter, liaison and analyst. The leader directed the group's work, leading the discussion and ensuring that everyone participated. The reporter recorded the group's answers to the problems and submitted the final work. All students in the group were required to approve and sign the submitted work. The liaison was the person who was sent to get information from the instructor, the classroom computer or other resources as needed during the group discussion. The analyst noted the problem solving and learning process used by the group during the discussion. The analyst added a statement to the final report about the process used by the group to arrive at consensus answers when students did not agree about the correct answer to a problem or to answer questions that none of the group members had been able to answer independently. Each student in the class was required to assume each role at least twice during the semester. Group reports were graded, all students in the group earned the same grade for the group work (except those who had not submitted individual answers before class began). Group work constituted 12% of the students' final grade for the course.

Assessment and Analysis

Exam questions for this class always include items that test the students problem-solving and analytical skills as well as their knowledge and comprehension of concepts. I do not return final exams to students and I purposefully reuse some exam questions for the purpose of assessing any changes I make to my class. I identified assessment items that were identical in the exams from the first two years of the study and included these on the exams for the second two years of the study. The questions used for assessment of the project were classified as knowledge, comprehension or analysis/application as described by Bloom (1956). The assignment of assessment items to Bloom categories was verified by two colleagues, who teach other sections of the same course. Seven knowledge questions, eight comprehension questions and seven application and analysis questions were used in the analysis. Representative questions and grading rubrics are provided in Appendix A. To control for grading bias, a colleague who teaches a different section of the course independently graded the questions shown in Appendix A using the rubric provided. This was done for 10 randomly chosen tests from each experimental group (group work, no group work). To compare experimental groups, I calculated an average score for responses to all of the questions in each Bloom category for each student. I used a two-tailed Student's t-test statistic to compare the average scores of students taught using the class format with no group work (2005 and 2006) to the average scores of students who participated in the group problem-solving sessions (2007 and 2008). An analysis of covariance was also conducted using SAT/ACT scores as a covariant for each student's average score for each category of exam questions. The statistical analysis was conducted using SSPS programs.

Institutional Review Board Approval

Approval for the analysis and publication of the results of this study (exempt status) were obtained from the Hofstra University IRB.

Results

General Observations

The first indication of a general improvement in the student learning when group problem-solving sessions were incorporated into the course (Fall 2007 and 2008) was improvement in mid-semester test grades. This continued throughout the semester such that the final overall average grade in the class was higher in the years when students worked in groups (Table 1). The change in the class grade average may have been due to two things unrelated to true improvement in learning. First of all, in 2007 and 2008 group work and homework (preparation for group work) constituted 18% of the final grade whereas in 2005 and 2006 homework constituted only 16% homework of the student's grades. However, the improvement in performance is still apparent in the average scores on in-class tests. The average score (%) on three mid-term and the final exam increased steadily throughout the period of the study. Average score on final exam scores alone also improved with each consecutive year of the study (Table 1). The difference in test scores, however, was confounded by the fact that in 2007 and 2008 the students had the benefit of additional practice. In 2005 and 2006, the students practiced problems on their own and were given feedback. In 2007 and 2008, students practiced the problem first on their own and then in a group before getting feedback. In addition, average combined Verbal and Math SAT vary slightly from year to year. However, the pattern of changes in SAT does not correlate with the changes in final grade and test scores (Table 1).

Table 1. Comparison of learning gains for students who did not participate in group problem-solving session (2005 and 2006) and those who did participate in group problem-solving sessions (2007 and 2008). Test scores are the average for all tests taken by all students in a class expressed as a percentage. Combined SAT is mean Verbal plus Math SAT scores for all students in the class.

Year	N	Final grade	Test Scores (%)	Final exam (%)	Combined SAT
2005	38	71.4	55.9	53.0	1138
2006	35	67.6	58.6	58.2	1155
2007	41	76.2	62.4	64.5	1112
2008	41	77.1	70.2	65.4	1169

<u>Improvements in final exam scores</u>

As a more rigorous evaluation of the effectiveness of the group problem-solving sessions in improving student learning, I compared the average score on identical final exam knowledge, comprehension and application/analysis questions obtained by students in the two classes when no group work was done to the average scores obtained by students in the two classes that incorporated group work. Data from 2005 and 2006 were combined and categorized as "no group work" format, which included lecture plus assignments with no group problem-solving workshops. Data from Fall 2007 and 2008 were combined and represented the "group work" format in which group problem-solving sessions were added. This analysis was done separately for each category (knowledge, comprehension and application/analysis) of exam question. The overall mean final exam score for all students who had participated in group-problem-solving sessions was higher in all three categories than the overall mean scores for students who had not participated in group work (Table 2). In addition, there was a significant difference in the average scores for each assessment category under the two class formats ($p < 0.05$ using two-tailed Student's t-test). The difference in average knowledge and comprehension scores were highly significant.

Table 2. Comparison of final exam knowledge, comprehension and application and analysis scores between student who did and did not participate in group problem-solving sessions. P value based on two-tailed Student's t-test.

Category	p value t-test	Teaching Format	Mean	N	Std dev	Std Error Mean
Knowledge	1.97×10^{-6}	No group work	56.784	73	20.32	2.349
		Group work	72.943	82	19.94	2.216
Comprehension	6.15×10^{-3}	No group work	40.629	73	16.42	1.922
		Group work	48.378	82	18.11	2.000
Application/Analysis	1.0×10^{-2}	No group work	35.058	73	14.93	1.747
		Group work	42.360	82	20.15	2.225

Improvement in exam scores after adjustment for SAT scores

Because there was a difference in the average SAT/ACT scores between the students in each semester used in this study, an analysis of covariance (ANCOVA) was conducted using SAT/ACT scores as a covariant with each category of test scores as the dependent variable. As above, data from 2005 and 2006 were combined and categorized as "no group work" format. Data from Fall 2007 and 2008 were combined and represented the "group work" format. The independent variable was the course format. A significant increase in exam scores was seen in each category when the "group work" format was employed (Figure 1).

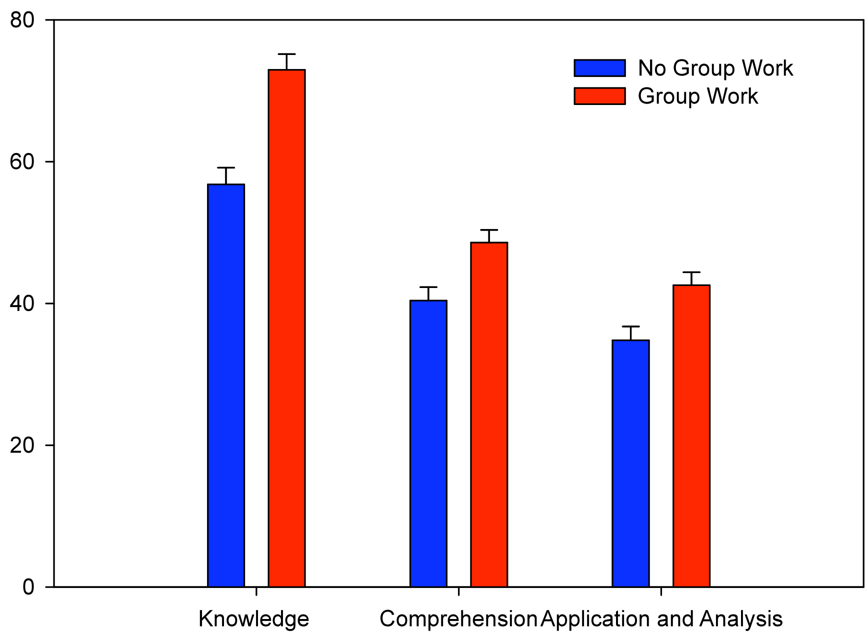

Figure 1. Mean scores for knowledge, comprehension and application and analysis questions corrected for variance in SAT/ACT scores. Error bars represent standard errors.

Knowledge Score Results

The uncorrected mean score on knowledge questions improved greatly with the introduction of group work to the course format (Table 2). Differences in SAT score between students in the two course format groups did not have an influence on scores for knowledge questions (F 1,152 = 1.86, p = 0.175). Course format itself did have a significant effect on knowledge scores (F 1, 152 = 25.0, p =.000). Because SAT scores were not a significant covariate there is no difference between raw and corrected mean scores for knowledge questions between the two course format groups. The mean score on knowledge questions was more than16 points higher when group work was added to the course format (Table 2).

Comprehension Score Results

The uncorrected mean score on comprehension questions also improved with the introduction of group work to the course format (Table 3). In this case, however, SAT scores have a significant effect on scores for comprehension questions (F 1,152 = 22.99, p = .000). After adjusting for SAT scores as a covariate, course format had a significant effect on scores for comprehension questions (F 1,152 = 22.99, p = .002). The corrected mean scores on comprehension questions was 8 points higher when group work was added to the course format (Table 3).

Table 3. Comprehension mean scores for students under different course formats. Uncorrected values and values corrected for SAT covariate are shown.

Course Format	**Uncorrected scores**		**Corrected scores**		**N**
	Mean	**SD**	**Mean**	**SE**	
No Group Work	40.6288	16.41997	40.409(a)	1.898	73
Group Work	48.3780	18.10963	48.573(a)	1.790	82

(a) Covariates appearing in the model are evaluated at the following values: SAT Comb = 1142.9677.

Application and Analysis Results

The uncorrected mean score on application and analysis questions also improved with the introduction of group work (Table 4). Again, SAT scores have a significant effect on students' ability to apply knowledge and to analyze information (F 1,152 = 28.33, p = .000). When adjusted for SAT scores as a covariate, course format still has an effect on the student scores for application and analysis questions (F 1,152 = 22.99, p = .004) and the corrected mean score on application and analysis questions was also nearly 8 points higher when group work was added to the course format (Table 4).

Table 4. Application and Analysis mean scores for students under different course formats. Uncorrected values and values corrected for SAT covariate are shown.

Course Format	**Uncorrected scores**		**Corrected scores**		**N**
	Mean	**SD**	**Mean**	**SE**	
No Group Work	35.0575	14.92884	34.810(a)	1.928	73
Group Work	42.3598	20.14709	42.580(a)	1.819	82

(a) Covariates appearing in the model are evaluated at the following values: SAT Comb = 1142.9677.

Discussion

This project examined the impact of group problem-solving sessions on learning in an introductory cell biology and genetics course. Mean scores on a set of final exam knowledge, comprehension and application and analysis questions for students who participated in the group sessions improved significantly when corrected for variability in SAT scores over those of students in the same course taught without the inclusion of the problem-solving sessions. The class under both teaching formats included only freshmen and was taught at the same time on the same days of the week, in the same classroom with the same instructor using the same textbook and many of the same teaching materials, including key questions, problems and case study materials. The only parameters that changed in any noticeable way were 1) the particular set of students, which can have profound effects on learning outcomes but is not controllable and 2) the substitution of group problem-solving workshops in lieu of lecture for about 30% of the class lecture sessions.

It is particularly interesting and encouraging that the mean scores on the knowledge questions improved so dramatically. One criticism of assessments of this type of teaching intervention is that documented improvements may be due only to practice effects. One group of students is given the opportunity to practice a particular skill and the students with whom they are compared were not given this opportunity. Both groups of students are then assessed on their ability to perform the skill in question and, of course, those who have practiced this skill outperform the control group. This criticism can be applied to the results of this study with the exception of the improvements in scores on knowledge questions. Knowledge questions probed the basic content knowledge of the students. If anything, the students in the "group work" teaching method group had less opportunity to acquire basic knowledge, at least in the classroom setting, because basic knowledge was presented for approximately 30% less time than under the "no group work" teaching format. Nonetheless, the students who were taught using the new format performed significantly better on basic knowledge questions on the final exam. This result is encouraging for a number of reasons. First of all, it supports the idea that the substitution of student-centered learning for lectures does not result in the acquisition of less knowledge. In fact, in this case, less lecturing resulted in more knowledge acquisition. Secondly, the improvement in scores on knowledge assessment items, when these items were not a part of what was practiced, suggests that the improvements in comprehension and application and analysis assessment items may also not be due only to a practice effect.

Similar results with cooperative or group learning have been reported in the past. A meta-analysis of outcomes from different cooperative learning techniques used at a variety of grade levels from middle school through post-baccalaureate indicated that cooperative learning results in greater student learning and more positive attitudes regardless of the particular technique used, although some techniques seemed to be more efficient than others (Johnson et al. 2000). A different meta-analysis that focused on undergraduate science, math and pre-health sciences courses showed significant positive effects of small group learning formats on achievement, persistence and attitude (Springer et al. 1999). The effect on achievement was greater for students in four-year as opposed to two-year colleges. It was also greater for groups reported as exclusively or predominately African American and Latino in comparison to mixed ethnic groups. In this meta-analysis the effect of all three parameters was greater when the investigator was the instructor (as in this study), suggesting investigator bias. The effects of small-group learning on achievement were also significantly greater when measured with instructor-made exams (as in this study) than with standardized instruments. The positive effects on achievement were greater for out-of-class meeting than for in-class group work. Published examples of the power of cooperative and group learning can be found for all fields of science, mathematics and technology (Hake et al. 1998, Crouch 2001, Born et al. 2002, Tenney and Houch 2003, Jones-Wilson 2005, Knight and Wood 2005, Carmichael, 2009) It is rare for an investigator to report negative or no effects (but see Dochy et al., 2003).

Two biases may confound the results reported here: 1) grading bias, which can be controlled and 2) subtle changes in the way material is presented in lecture, which cannot be controlled unless scripted lectures are used. For this study, the author is both the investigator and the instructor. It is possible that scoring on open response questions on exams was unconsciously biased in favor of the students in the “group work” cohort. However, grading rubrics, which were created before the beginning of the study, were used to grade open response question (see Appendix A for examples). In addition, a colleague who teaches a different section of the same course used my rubric to re-graded the responses to the questions shown in the Appendix for ten students from each comparison group in the study. While there were very small differences in some individual scores, the mean scores for the comprehension and application/analysis questions did not change. The other possible bias is much more difficult to control. I learned from year to year which content was most difficult for the students and made slight changes in my presentation of difficult materials based on this knowledge. It is possible that these changes or small changes in emphasis contributed to the reported improvements. There is no post-hoc method to compensate for this possible bias. Another criticism of this study is the small number of students involved in the study. It is possible that the improvements that were seen have nothing to do with the teaching method but are an artifact of a particular, highly motivated and cooperative group of students or some other uncontrolled variable.

It will be interesting to see if the short term changes in knowledge, comprehension and ability to apply and analyze genetics and cell biology concept is retained over a more prolonged period of time. I plan to use the results of the cell biology and genetics components of the ETS Biology majors field test to examine long-term retention of learning. All biology majors at the university are required to take this comprehensive biology subject test in their final semester. The test assesses only content knowledge and is not a good indicator of students’ critical thinking skills. However, in four years I will be able to compare the ETS/Biology test scores of students in my section of the introductory cell biology and genetics course who did and did not participate in class problem-solving sessions. I will also be able to compare my students’ scores to those of all other biology majors.

Acknowledgements

This work was supported by NSF-EHR #0314910.

References

Born, W.B., Revelle W. and Pinto, L.H. (2002). Improving biology performance in workshop groups. Journal of Science Education and Technology 11 (4): 347-365.

Boyer Commission on Educating Undergraduate in the Research University for the Carniege Foundation for the Advancement of Teaching. (1998). Reinventing Undergraduate Education: A Blueprint for America's Research Universities. Available at http://naples.cc.sunysb.edu/Pres/boyeer.nsf.

Carmichael, J. (2009). Team-based learning enhances performance in introductory biology. Journal of College Science Teaching 58(4): 54-61.

Crouch, C.H. and Mazur, E. (2001). Peer Instruction: Ten years of experience and results. American Journal of Physics 69 (9): 970-977.

De Hann, R.L. ((2005). The impending revolution in undergraduate science education. Journal of Science Education and Technology 14(2): 253-268.

Dochy, F.M., Segers, M., Van den Bossche, P. and Gijbels, D. (2003). Learning and Instruction 13: 533-568.

Hake, R. (1998). Interactive-engagement vs. traditional methods: A six-thousand student survey of mechanics test data for introductory physics courses. American Journal of Physics 66: 64-90.

Handelsman, J.D., Miller S. and Pfund C. (2007). Scientific Teaching. NewYork: W.H. Freeman and Company.

Johnson D. W., Johnson R. T., Stanne M. E. (2000). Cooperative Learning Methods: A Meta-Analysis. Cooperative Learning Center website (www.clcrc.com).

Jones-Wilson, T.M. (2005). Teaching Problem-Solving Skills without Sacrificing Course Content. Journal of College Science Teaching 35(1): 42-46

Knight, J.K. And W.B. Wood 2005. Teaching More by Lecturing less. Cell Biology Education 4:298-310.

Mazur, E. (1997). Peer Instruction: A User's Manual. Upper Saddle River, NJ: Prentice Hall, Inc.

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

National Research Council (NRC). (2000). How People Learn: Brain, Mind, Experience and School. Washington, DC: National Academy Press.

National Research Council (NRC). (2003). BIO 2010: Transforming Undergraduate Education for Future Research Scientists. Washington, DC: National Academy Press.

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

Springer, L., Stanne, M.E. and Donovan, S.S. (1999). Effects of small-group learning on undergraduates in science, mathematics, engineering and technology: A meta-analysis. Review of Educational Research 69: 21-50.

Tenney, A. and Houck, B. (2003). Peer-led team learning in introductory biology and chemistry courses: A parallel approach. Journal of Mathematics and Science: Collaborative Explorations 6: 11-20.