Chemical Education Research

edited by
Diane M. Bunce
The Catholic University of America
Washington, D.C. 20064

Group Problem-Solving versus Lecture in College-Level Quantitative Analysis: The Good, the Bad, and the Ugly

Vickie M. Williamson* and Marvin W. Rowe
Department of Chemistry, Texas A & M University, College Station, TX 77843-3255; *williamson@tamu.edu

Significance

Students often become disinterested in chemistry when the course involves only their passive note taking (*1*). An increase in positive attitudes often results when there is a move away from lecture, toward a group-work format (e.g., *2, 3*). This study originated from discussions between the authors, a veteran instructor who had more than 30 years of experience (MWR), and a chemical educator (VMW). Although the instructor was not aware of educational literature indicating that cooperative learning can lead to statistically significant improvement in academic achievement (*4–12*), he was convinced that the lecture mode was not achieving the aims of the course and sought a better means of achieving those ends. His main goals were that the students learn to think about the problems they encounter in quantitative analysis, rather than to memorize solutions, and willingly seek assistance when their thinking does not lead to the problem solution. Over the last 15 years, the instructor had tried a number of techniques involving cooperative learning, but wanted more than anecdotal evidence for his decision to abandon traditional lecture. The study looks at a complete replacement of lecture with group problem-solving sessions in a junior-level quantitative analysis class for non-chemistry majors.

Objectives of the Study

The following research questions were investigated.

1. Do students who attend group problem-solving sessions in place of lecture in quantitative analysis learn the content as well as students in a traditional class?
2. Do the attitudes and dropout rate of quantitative analysis students who attend group problem-solving sessions differ from those of students who attend traditional lecture?

Subjects

The subjects were students enrolled in two sections of a junior-level quantitative analysis course at a major southwestern university. Both sections were taught by MWR. The sections met on the same days of the week at slightly different times. One section was randomly selected to be the treatment section.

Treatment

The treatment consisted of replacing a traditional lecture with cooperative group problem-solving throughout the semester. The control section received traditional lectures on the content, during which the instructor worked sample problems and explained the theory behind each step. Both the treatment and control sections had reading assignments to complete before each class period.

In the treatment section, students were assigned to heterogeneous cooperative groups of four (to encompass a variation in TOLT scores and in gender) and were presented problems on an overhead projector or in a handout. The students were asked to solve the problems with full discussion within their group. After one member of the group solved the problem, the group was required to discuss the problem until all members understood it. After all the groups had finished, a member of the group was randomly selected to begin the solution on the board or the overhead projector, or to direct the instructor through the steps to solve the problem. Sometimes members of a single group completed the solution; other times a member of one group did one part of the solution and a member of another group finished the problem. The rest of the class was encouraged to find flaws and patterns in the reasoning. Problems used in the control and treatment sections were equivalent in content and number. In the event that the treatment section didn't finish all the problems in one class period, they were asked to finish them as homework before the next class.

Data Collected

Both qualitative and quantitative data were collected. The subjects were given the Test of Logical Thinking (TOLT) on the first day of class. The TOLT measures the following reasoning abilities: controlling variables, proportional reasoning, combinatorial reasoning, probabilistic reasoning, and correlational reasoning (*13*). TOLT scores can range from 0 to 10. The scores from the TOLT were used to construct the heterogeneous student groups.

To measure content knowledge, exams, quizzes, and the course average were used. Both sections took three common exams given at night: exam 1, exam 2, and the final exam. To ensure that students were prepared for the daily classroom discussions, frequent "daily" quizzes on the reading assignments and on classroom activities were given to each section. Quiz questions were similar to those on the hour examinations, but were graded less rigorously. Course averages were calculated using the formula [E1 + E2 + (2 × Final) + (0.5 × sum of quizzes)]/4.5.

A 7-point Likert-scaled attitude survey similar to those used in other research studies (*14, 15*) was constructed. The seven-point scale permitted the desired degree of variation in responses. This 31-item instrument was used at the end of the semester. It asked students about their perception of how they understood the chemistry content, their attitude toward chemistry, and their attitude toward the teaching

Table 1. Content Understanding and TOLT Scores

Group	Exam 1		Exam 2		Final Exam		Quizzes		Overall Course		TOLT	
	Av (SD)	*p*	Av (SD)	*p*	Av (SD)	*p*	Av (SD)	*p*	Av (SD)	*p*	Score (SD)	*p*
Control	72.58 (16.78)	.69	77.31 (17.95)	.88	73.92 (22.38)	.31	18.65 (17.24)	.54	74.38 (17.24)	.53	8.46 (1.45)	.81
Treatment	75.30 (22.28)		76.41 (18.81)		79.22 (12.29)		19.43 (12.29)		77.53 (3.55)		8.59 (1.70)	

methods used in their class. (See sample items in the results section.) Additionally, students were asked to write about what they liked most and least about the course. The full attitudinal questionnaire can be found online.[W]

VMW observed both sections and made field notes concerning student behaviors, interactions, etc. The treatment-section observations comprised notes on the class as a whole and on two specific groups. Notes about the groups included dialogue and interaction of group members. Interviews held during the semester with various group members consisted of both open-ended and leading questions, which are described later. Students completed an open-response midterm evaluation. MWR also kept notes about his impressions throughout the progress of the course.

Only students who completed all exams, the midterm evaluation, the end-of-term attitude survey, and the TOLT were included in the quantitative analysis. This restriction left 13 of the 27 students in the control section and 32 of the 52 in the treatment section. Other reasons for the larger number of subjects in the treatment section will be discussed later.

Results and Discussion

No content measure (quizzes, exams 1 and 2, the final exam, and the course average) using a multiple analysis of variance showed significant differences between sections at the $p < .05$ level when the students from whom all data were collected were considered (Table 1). There was also no significant difference in the TOLT scores of the sections ($p = .81$). Since the reasoning ability of the sections did not differ, we treated them as similar groups. Even using the TOLT score as a covariant showed no differences in the content scores between the sections when reasoning ability was controlled.

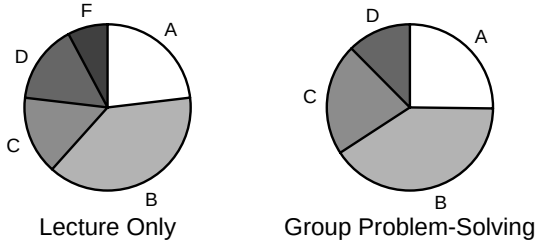

Figure 1. Letter grades for those finishing the course.

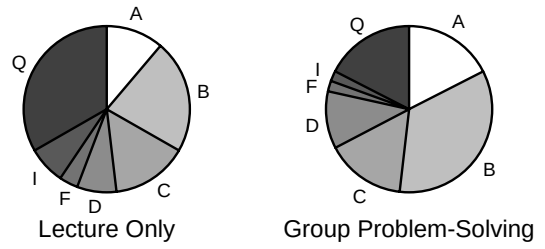

Figure 2. Letter grades for all students beginning the course. I is incomplete; Q is Q–drop withdrawal.

The course average was 74.38% (SD 17.24) for the lecture section and 77.53% (SD 3.55) for the treatment section; the difference was not significant. Grade cutoffs were the same for both sections: A = 89; B = 76; C = 65; D = 45; F < 45. Figure 1 shows the letter grades for the subjects who completed the course and were included in the original study. The proportion of A's, B's, and C's was 76.9% in the control section ($n = 13$) and 87.5% in the treatment section ($n = 32$).

A more interesting finding is the grades for all those who enrolled in the course (27 students in the control and 52 in the treatment section) (Fig. 2). The proportion of A's, B's, and C's was 48.1% of those enrolled in the control section and 67.3% of those enrolled in the treatment section. The withdrawal rates were very different for the two sections (33.3% for the control section and 17.3% for the treatment). It should be noted that the withdrawal system of "Q–drop" allows students only 3 such drops in their academic careers, but a Q–drop also removes all mention of the course from a transcript. Students hoard their Q–drops and are reluctant to use them, except in urgent cases. A χ^2 analysis of the Q and I scores versus the A–F grades for each section showed a significant difference between the sections ($p = .040$). The rationale for combining the Q and I rates is that the I's are students who quit participating in the class, but did not officially drop. Since most institutions report grade distributions as percentages, the difference between a 33.3% and a 17.3% withdrawal rate may be a better way of reporting this finding.

Two of the 31 items on the attitude survey showed significant differences at the $p < .05$ level when analyzed with a multiple analysis of variance (see Table 2). The findings on the first item might reflect a desire of the control section to discuss more, since there was little, if any, discussion in this section. For the treatment section, the mean on this item is skewed toward a major component of their treatment, learning by listening to another student's viewpoint. The means for the second item reflect the method by which students were asked to learn in this study. While in some ways this finding isn't surprising, the factors in the difference between the two sections should be investigated. Were the students who finished the course those who preferred to learn by the method used in their class?

Items on which the two sections were similar also were informative (see Table 3). Both sections felt that they understood the words used, were able to solve problems, and learned a lot when they worked problems from the end of the chapter. Although neither section felt comfortable discussing things with the people in the class, interviews revealed that students thought this item meant before the *full* class, not working in groups to solve problems. It is troubling that both sections thought there was only one way to work a problem.

Two groups from the treatment section were observed in detail. The more notable group in terms of ability to verbalize their mode of operation will be described here. Observations

Table 2. Significant Differences on Attitude Survey

Item, Scored on Likert-Type Scale of…			Group Mean Score (SD)		*p*
1	to	7	Control	Treatment	
I learn more when I hear another student's viewpoint	…	I don't learn more by hearing the viewpoint of another student	2.54 (1.27)	3.59 (1.27)	.0292
I learn more by listening to the lecture	…	I learn more by working the problems	3.59 (1.60)	5.44 (1.56)	.0016

Table 3. Similiarities on Attitude Survey

Item, Scored on Likert-Type Scale of…			Group Mean Score (SD)		*p*
1	to	7	Control	Treatment	
I don't understand the words used	…	I understand the words used	6.18 (0.60)	6.00 (0.89)	.54
I can't solve the problems	…	I am able to solve the problems	3.59 (1.60)	5.44 (1.56)	.67
I learn a lot when I work problems at the end of the chapter	…	I learn nothing when I work problems at the end of the chapter	2.00 (0.77)	2.29 (1.24)	.47
I feel comfortable discussing things with the people in my chemistry class	…	I don't feel comfortable discussing things with the people in my chemistry class	5.18 (0.87)	5.32 (1.66)	.79
I enjoy working in a group to solve problems	…	I don't enjoy working in a group to solve problems	3.00 (1.61)	3.10 (1.81)	.88
The methods other students use to work a problem may differ from mine	…	There is only one way to work a particular problem	6.00 (1.00)	5.39 (1.56)	.23

of the second group and of the class as a whole coincided with those of the notable group. These observations offered insight into the atmosphere in the treatment section.

The members of this group were Tom, Ken, Natalie, and Joy. Tom was frequently the first to finish the problem. Ken and Joy seemed to have good ability and often worked together first. Tom would then confer briefly with Ken and Joy. When these three agreed, Tom would work with Natalie while Ken and Joy worked together to finish their own understanding. Ken and Joy joined the conversation between Tom and Natalie to help or to gain clarification. While Natalie was not "clueless", she was definitely the weak one in the group.

Observations were made of the group interactions during the class, and group members were interviewed individually at various times during the semester. In the interviews, they were asked open-ended questions about their perceptions of the particular problems or lesson that day and their perceptions of how the group was functioning. Depending on the responses, they were then asked leading questions, such as "Do students with lower abilities hold you back?" Tom and Natalie provided the most interesting views (see the assertions).

Analysis of qualitative data lends support to a number of assertions concerning the effects of the treatment. Qualitative data included observation notes, lecturer's notes, the open-response comments from the midterm evaluation, the open-response comments from the end-of-semester attitude survey, and interviews with group members. Findings from all these sources were categorized and collapsed into major areas. These areas were compared to the quantitative data. Triangulation of data sources was used to develop assertions.

For example, the process that led to assertion 3 included these components. The first author's general observations and field notes of the control section painted a very austere picture. The students sat quietly without talking before, during, or after class. They took notes but asked very few questions, even when the professor asked for questions. The same data for the treatment section were very different. The groups of four students sat together. The group members talked to each other, even before class. The atmosphere was much more casual. When the groups were presented with a problem, work and discussion began. Most often the students first tried to work the problem on their own and then they discussed it with the group.

The instructor described the same differences between the sections. The more important differences that the instructor observed in the treatment section were (i) much better rapport between the instructor and the students, (ii) reduced reluctance to ask questions in class, and (iii) more frequent visits to the instructor's office to ask questions. These findings, plus the open-response questions on the attitude instrument and the interview data, led to the third assertion. Sample excerpts from interview data and open comments from the attitude survey follow many of the assertions.

The Good

Assertion 1. Even bright students are more self-assured when they have more opportunities to exercise/verbalize their understandings and abilities.

QUESTION: Do students with lower abilities hold you back?

TOM: No, being the teacher makes you really, really learn it.

Assertion 2. It seems that students of slower ability can be brought "up to speed" more quickly by a peer, often owing to the reluctance to see the instructor. Students are less intimidated when a peer points out errors. (More intimidated students withdraw from the class.)

QUESTION: How do you feel about your group teaching you?

NATALIE: I would have dropped if they didn't help me.

QUESTION: Couldn't you go to the professor?

NATALIE: No, my questions are too dumb. My group can help me understand.

Natalie's comment seems to support this assertion, as does the interesting fact that the numbers of students who withdrew from the course differed for the sections. Only 17.3% of the treatment section withdrew, whereas 33.3% of the control section withdrew. This difference was found even though the two sections had similar reasoning ability (TOLT) scores,

whether those completing the course, those enrolled in the course, or those withdrawing from the course were considered. While the students withdrawing from the course had a somewhat lower TOLT average than those finishing the course, the scores for the two sections of withdrawals were similar (7.50, SD 1.31, for the control and 7.67, SD 2.88, for the treatment).

Assertion 3. The feelings of comradeship will enable students to persist, whereas feelings of isolation lead to withdrawals. Students in the control section had the perception that they were in competition with other students in their class. This is different for the students in the treatment section, who worked together at least within their group. Natalie's comments above are typical of the comradeship felt by the groups in the treatment section. Likewise, open-response comments to survey data supported the notion that the control section felt they were competing with each other and working in isolation, while the treatment section felt that they were capable of succeeding in this difficult class by struggling together with the instructor. In addition, although students in general are reluctant to seek an instructor's help, students in the control section did not ask questions in class or visit the instructor's office with questions at the same rate as students in the treatment section. The instructor noted about one-half the rate of office visits and an even lower rate of in-class questions in the control section. A student from the control section, upon noting that students from the treatment section were visiting with one another before one of the common exams that were given in the evening, commented to the instructor that the treatment section was more closely knit than his own section.

The Bad and Ugly

Assertion 4. Students and instructors resist new methods. In the treatment section, 28.1% of the students said they would have liked to have some lectures, but over half of these also said that working problems or group work helped them understand. Implementation of lecture replacements will not be favored by all. While the instructor in this study was a proponent of cooperative groupings, some instructors resist new methods. Part of the motivation for the study was to provide evidence for colleagues who were interested, but hesitant to change their own methods, and for those who were critical of the new methods.

Assertion 5. Group problem-solving as the sole approach has drawbacks. This study was carried out through the complete semester. Observations of the groups and their comments on the attitude survey showed that many groups tired of a repeated routine. Perhaps the implementation of group problem-solving as *one* of the methodologies to improve student attitudes and persistence in the course would be more fruitful. Further investigation is needed into whether a combination of methods might be the best answer for the instructor. Wright reported success with cooperative methods and open-ended labs (*12*). It may be that each academic environment must search out the best combination of techniques its instructors can offer to serve the needs of its own student population.

Conclusions

Group problem-solving sessions appear to be viable replacements for traditional lectures in courses such as quantitative analysis where the "lectures" usually amount to watching the instructor, the expert, work problems for the novices. In this study, students in the treatment section perceived that they understood more by working problems, were less likely to withdraw from the course, and had better feelings toward the whole educational experience. Similar findings on academic achievement have been reported (*1–12*). The use of group problem-solving seems to give students the support they need to persist in courses like the one in this study. The reduced number of withdrawals in the group-problem-solving section (just over one-half the rate in the control section) deserves particular attention, especially as we try to attract and maintain more students in science majors. Group problem-solving should be in our arsenal of teaching strategies.

Epilogue

Armed with findings from other literature and from this study, the instructor now disavows traditional lecturing, in favor of the group problem-solving method described in this paper. However, he does give short lectures on occasion. In his short "lectures", he makes a few intentional mistakes as he solves problems out loud and encourages the students to find fault with his solutions or thought process. This is an effort to encourage discussion with groups and within the class as a whole and to give a change in the class routine. Group problem-solving is a mainstay of his course.

[W]Supplemental Material

The complete attitudinal questionnaire is available in this issue of *JCE Online*.

Literature Cited

1. Bodner, G. M. *J. Chem. Educ.* **1992,** *69*, 186–190.
2. Gosser, D. K.; Roth, V. *J. Chem. Educ.* **1998,** *75*, 185–187.
3. Ross, M. R.; Fulton, R. B. *J. Chem. Educ.* **1994,** *71*, 141–143.
4. Johnson, D. W.; Johnson, R. T. *Cooperation and Competition: Theory and Research;* Interaction Book Company: Edina, MN, 1989.
5. Blosser, P. E. The Science Outlook: *Using Cooperative Learning in Science Education;* ERIC-CSMEE: Columbus, OH, 1993; pp 1–9.
6. Johnson, D. W.; Johnson, R. T.; Smith, K. A. *Cooperative Learning: Increasing College Faculty Instructional Productivity;* ASHE-ERIC Higher Education Report No. 4; George Washington University: Washington, DC, 1991.
7. Slavin, R. E. *Educ. Leadership* **1991,** *48*, 71–82.
8. Walters, J. P. *Anal. Chem.* **1991,** *63*, 977A–985A.
9. Walters, J. P. *Anal. Chem.* **1991,** *63*, 1179A–1191A.
10. Wenzel, T. J. *Anal. Chem.* **1995,** *67*, 470A–475A.
11. Wenzel, T. J. *Anal. Chem.* **1998,** *70*, 790A–795A.
12. Wright, J. C. *J. Chem. Educ.* **1996,** *73*, 827–832.
13. Tobin, K.; Capie, W. *Educ. Psychol. Meas.* **1981,** *41*, 413–423.
14. Abraham, M. R.; Williamson, V. M. In *Successful Approaches to Teaching Introductory Science Courses;* McIntosh, W. J.; Caprio, M. W., Eds.; Society for College Science Teachers: Cedar City, UT, 1992; pp 21–28.
15. Abraham, M. R.; Renner, J. W. *Sequencing Language and Activities in Teaching High School Chemistry: a report to the National Science Foundation;* University of Oklahoma: Norman, 1983; ERIC Document ED 241 267.