DOCUMENT RESUME

ED 412 099	SE 060 616
AUTHOR	Wright, John C.; Woods, R. Claude; Millar, Susan B.; Koscuik, Steve A.; Penberthy, Debra L.; Williams, Paul H.; Wampold, Bruce E.
TITLE	A Novel Comparative Assessment of Two Learning Strategies in a Freshman Chemistry Course for Science and Engineering Majors.
INSTITUTION	Wisconsin Univ., Madison. LEAD Center.
SPONS AGENCY	National Science Foundation, Arlington, VA.; Advanced Research Projects Agency (DOD), Washington, DC.
PUB DATE	1997-01-00
NOTE	21p.
CONTRACT	DUE-9455928; EEC-8721545
PUB TYPE	Reports - Evaluative (142)
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	*Active Learning; *Chemistry; *College Science; Course Content; *Curriculum Evaluation; Educational Innovation; Educational Strategies; Engineering Education; Higher Education; *Majors (Students); Science Curriculum; Surveys; Undergraduate Study
IDENTIFIERS	University of Wisconsin Madison

ABSTRACT

This paper reports on a study of a structured active learning approach to an analytical chemistry course. The purpose of the study was to assess the longitudinal outcomes of this approach and to provide feedback to curriculum developers about the effects of the changes being made in the chemistry curriculum. The data, collected from a cohort of students in 1995, consists of open-ended interviews with students, open-ended student survey questions, Likert scale student survey items, and faculty and teaching assistant interviews. The results indicate that there is a number of longitudinal effects of structured active learning and although it is possible to assert that the comparison group does not report the kinds of learning outcomes that the structured active learning students report, further study is needed to be sure that the findings are reliable. Contains 11 references. (DDR)

A NOVEL COMPARATIVE ASSESSMENT OF TWO LEARNING STRATEGIES IN A FRESHMAN CHEMISTRY COURSE FOR SCIENCE AND ENGINEERING MAJORS

John C. Wright[1], R. Claude Woods[1], Susan B. Millar[2], Steve A. Koscuik[2], Debra L. Pen berthy[2], Paul H. Williams[3], and Bruce E. Wampold[4]
[1]Department of Chemistry
[2]LEAD Center
[3]Department of Plant Pathology
[4]Department of Educational Psychology
University of Wisconsin, Madison, WI 53706

ABSTRACT

In order to perform a credible evaluation of whether structured active learning methods result in significant improvements in student learning, the students in a section of a course with active learning were compared to those in a control section using oral interviews that tested student competence. Qualitative research methods were also employed to identify the reasons for any differences. The results show substantial differences in the students' reasoning and self expression skills that are directly attributable to their structured active learning experiences.

Introduction

Instructional methods in undergraduate science courses are currently the subject of intense interest, and a number of new approaches are being tried (1-8). A structured active learning (SAL) approach has been implemented over several years in an accelerated analytical chemistry course at Wisconsin (Chemistry 110) for well-prepared freshmen. The course is characterized by interactive classroom settings, cooperative student assignments and examinations, and somewhat open-ended group projects and laboratory experiments (9). Subjective evaluations of this experiment by both instructor and students had deemed it quite successful, consistent with experience of other instructors employing similar methods (1,7,8). The perception of those involved has been that structured active learning methods can improve enthusiasm and motivation and lead to improved thinking and reasoning skills (2-6). More objective or quantitative assessment of such methods, or comparison of students participating in SAL to those in control groups, is highly desirable, but has not been easily obtained. There are no agreed upon methodologies for making such a comparison.

In order to obtain a credible measure of the effects on student competence, faculty, staff, and students at Wisconsin cooperated in an experiment involving a multi-dimensional assessment strategy that was applied in the 1995 spring semester to the SAL section of analytical chemistry and a control section. Such evaluations involve complex and subtle issues of the relationships between course goals, assessment strategy, and the nature of what is being tested. The results are presented here not as a definitive judgment on the relative merits of the two educational approaches that were compared, but as an indication that it is possible to measure the extent to which the goals of a particular approach are in fact being achieved in an objective and quantitative way. These results do support the proposition that the students in the SAL section did quantifiably improve their reasoning and self expression skills.

Description of Experiment

The students enrolled from two sections of a first semester course in general chemistry (Chemistry 109). They are primarily science and engineering majors with advanced preparation that places them in the upper 10-15% of entering chemistry students. One section of Chem. 110 with 108 students was taught using cooperative learning methods and the other section with 95 students was taught using a more traditional approach with lectures and difficult homework problems. The faculty in both sections have strong reputations for teaching excellence, equivalent teaching evaluations, and teach the Chem. 110 course with similar material and depth.

There were, however, some significant differences between the sections in topics covered and problem solving approaches stressed. An *ad hoc* committee of chemistry faculty developed the guidelines that defined a credible assessment of the two sections. The resulting strategy employed 25 volunteer faculty assessors from external departments (10) both to conduct oral examinations with all students in both sections and to rank the students in competence. The students were divided into octiles based on their rank in class in Chem.109, and any faculty assessor typically saw 7-9 students within a single octile. The students were divided roughly equally between each section. Three faculty assessors were usually needed to cover all students within an octile. The strategy also involved extensive interviews of different student, faculty, and teaching assistant groupings by members of the Learning through Evaluation, Adaptation, and Dissemination (LEAD) Center at UW-Madison. The interviews were analyzed by the same qualitative research methods that are established for sociological research to identify the nature of the social interactions that characterize the learning environment and student performance (11, 12). The LEAD analysis of the course was done with freedom and independence from the course instructors. The LEAD analysis labeled the two sections as structured active learning (SAL) and responsive lecturing (RL) to correspond with their findings.

Results of LEAD Qualitative Assessment

The RL lecture section was characterized by a lecturing style that was effective in eliciting student responses and involvement in the lecture material. This approach also included very challenging homework problems, quantitative examination questions, well-defined laboratory experiments, and an open-ended laboratory project. Student cooperation was neither encouraged nor discouraged. The SAL lecture section used open-ended question-answer sessions during lectures and added cooperative learning methods that included an absolute grading scale, cooperative computer projects, open-ended laboratory projects, oral examinations on the project results, research paper analysis, and cooperative examinations. Both sections required the same textbook (13) but each used it primarily as a reference. The LEAD Center classroom observation and interview data established that the teaching strategy implementation for each lecture section was at the high end of the performance scale. Student attendance, attentiveness, and participation were very high in comparison with similar courses and students gave their lecturer high marks for the skill and care with which the lectures and course components were implemented. The interview data also established that although enrollment from Chem. 109 showed some difference in self-selection between sections, the differences did not correlate with the faculty assessor rankings. The average rank in class from Chem. 109 was 60.32 percentile for SAL and 60.75 percentile for RL, so students in both lecture sections had performed equivalently.

The interview data also showed sharp differences in the nature of the learning interactions. The SAL learning was characterized by student-student interactions. Students stressed the importance of the research-oriented, structured group activity and indicated that group interactions helped connect the lecture, laboratory, and other course components. Generally, students felt the class atmosphere fostered support and cooperation, although 20% of the students continued to work independently. The people who flourished in the cooperative environment enjoyed the challenge of solving open-ended problems and acquired greater self-reliance. They also felt that the course structure fostered an awareness of the complexity and frustrations that accompany genuine research. These observations show that the SAL section is an appropriate example of courses stressing active learning.

The RL learning was characterized by a focus on the lecturer as the authority for learning. There was an appreciation and respect for the professor with some viewing him as a role model. There was also a strong sense of accomplishment in mastering the material using the lecturer's step-by-step problem solving approach and mathematical modeling method. The professor was viewed as practically the sole source of information and understanding by some, while others mentioned the TAs and other students as valuable resources. Many students spontaneously formed groups that greatly assisted their learning, but other students preferred to work alone, or were unsuccessful in becoming a member of a suitable group. Some students reported frustration about their inability to connect lecture concepts and laboratory experiences. These observations suggest that the RL section is an appropriate example of courses stressing the lecturer as the focus for learning.

Results of External Faculty Assessment and Questionnaires

The faculty assessors were coordinated by an objective, external faculty member and the LEAD Center. Assessors constructed their own structures for the student orals, formed their own criteria for assessing competence, and filled out a pre-examination survey about their approach to the oral examination. They completed a second survey after the oral examinations that reported changes in their ideas, methods, and criteria for competence. Both faculty and students filled out questionnaires for each oral. LEAD Center personnel conducted follow-up interviews with all faculty assessors to document the nature of the orals. The faculty assessors were not told which lecture section was taught by SAL or RL, nor were the students' sections identified.

The questionnaire data are summarized in Table 1. The probability values or p values represent Mann-Whitney tests of significance (14) and can be considered to indicate the probability that the observed differences are in fact not different. The Mann-Whitney test was used because it does not depend on the nature of the data distribution. First, it is important to realize that the questionnaire data show that the students in both sections had very favorable answers to all the questions about satisfaction and accomplishment in the courses. This comfort with the course was reinforced by the LEAD interview data that showed students' feeling of accomplishment and satisfaction with the course was higher for both sections than other comparable courses. The faculty assessor data also show that the assessors were impressed with the knowledge and ability levels of the students in both sections. This observation reinforces the LEAD Center data that show both sections were taught at the high end of the performance scale. The questionnaire data show no differences in student nervousness between lecture sections both for the faculty and student questionnaires, but there are marked differences in all the questions on preparation and performance. Both student and faculty questionnaires showed very significant differences in the perception of the student preparation for future science courses. The student answers also showed substantial differences in the students' perception of how well they demonstrated their learning, how fluent they were in answering questions, and how knowledgeable they appeared. The SAL students appeared to spend 15% more time in out-of-class work and 56% more of that time working with other students, although the work load in both courses was substantial.

These differences in perception are also reflected in the performance. Assessors were asked to define both a relative rank where their students' competence was ranked from first to last (we label this approach "forced relative assessor ranking") and an absolute score where the individual students' performance was placed on a continuum from low to high competency (labeled "absolute assessor ranking"). The relative ranking strategy forces decisions that accentuate differences between students and make it easier to discern differences associated with teaching method. At the same time, it obscures the magnitude of the differences. It also introduces correlation between the two sections, because if one section does better, the other must do worse. This correlation prevents the use of statistical techniques based on the normal distribution. The absolute score gives complementary information about the magnitude of differences. An adjusted absolute rank was defined by grouping students with similar scores to distinguish them from students who differed more markedly. For example, if 4 students were clustered near the top of the continuum, 1 student scored above average, 2 students were clustered near the middle, and 1 students scored lower, the adjusted absolute rank for this assessor would have values from 1 (best) to 4. Thus, the number of values and the poorest value differed between faculty members but this ranking strategy eliminated the effects of out-lying values and much of the correlation problem. It could not be considered an absolute measure of student competence since the scores given by each assessor were clearly referenced to the students that assessor interviewed.

The differences in relative ranking between sections are the most significant indicators. The typical assessor had 4 students from each section, whom he/she ranked first though eighth. If all 4 students in one section were ranked ahead of the other 4, the average rankings would be 2.5 and 6.5 and their difference would be 4 for that assessor. When this approach was averaged over all 25 assessors, the maximum possible difference was actually 3.46 because some faculty gave some students equal ranks and some faculty saw different numbers of students. In order to see whether there were statistically significant differences between sections, the Wilcoxon matched pair signed-rank test was employed using the relative ranks of each assessor and the sign test was employed using the sign of the rank difference between sections (14). Both methods are robust, nonparametric statistical tests that are independent of the type of statistical distribution. They are the most common tests of significance when the data is not normally distributed. Table 1 shows the relative rank is 4.80 for the RL section and 3.68 for the SAL section. The 1.12 difference between sections is statistically significant with a p-value of 0.0066 for the Wilcoxon matched pair signed-rank test and 0.023 for the more rigorous sign test and it is 1/3 of the maximum possible difference that could have occurred. The section differences from each assessor are shown in Figure 1. The student octile seen for each assessor are indicated at the top of the figure. The largest differences are seen at the bottom and the middle octiles but the differences are significant for all octiles.

Similar results were found for the adjusted absolute rank of overall competence (see Table 1). The differences in sections for the adjusted absolute rank and the overall competence question were very significant at a much lower p value, because they represented independent observations. The relative rankings, the adjusted absolute rankings, the student competence question, and the grades the students assigned to themselves all reflect substantial differences in the competence demonstrated by the students.

Analysis of Correlations

In order to determine the nature of the individual assessor's exams and to identify further reasons behind the differences, the LEAD Center analyzed the faculty assessors' criteria for the relative rankings using the data from the personal interviews, the pre- and post-assessment surveys, the faculty reports on each student, and the student surveys about the oral exam. All the assessors asked students to demonstrate a basic knowledge of chemistry and an ability to use the knowledge in a way that required an integration of abstract principles. The differences in assessor approach were grouped into two broad categories labeled outcomes or process. Assessors in the outcomes group used the examination to measure the command of the material, while assessors in the process group used the examination to observe how the students approached new problems. The outcomes category was further subdivided into an analogy subgroup (6 of 25 assessors) that used problems requiring students to relate a new problem to their course material and an analysis subgroup (4 of 25) requiring students to solve a problem that was unrelated to the course material. The process category was also subdivided into an agility subgroup (4 of 25) that measured how rapidly students could analyze and solve a problem or how effectively they could react to new information and a meta-awareness subgroup (11 of 25) that focused on the thinking patterns (Did the students: self-correct, have a variety of perspectives, understand the larger context surrounding a particular problem, or relate theory and practice?). The primary criterion used by each assessor is indicated on the bottom of Figure 1.

The relative rankings showed strong correlations with the assessor subcategory. The results are summarized in Table 2. The assessors in the meta-awareness subgroup found the largest differences between sections, almost 1/2 the maximum possible differences. This finding indicates that the major reason for the large difference in student competence was the thinking process that the students displayed during the oral examination. The agility subgroup also found substantial differences. In the analysis subgroup, the differences became smaller and in the analogy subgroup, the differences nearly vanished. These results have two important implications. First, it is interesting that a large proportion of the faculty tested for meta-awareness as their primary criterion for competence, even though this criterion may not be what they test for in their own courses. Secondly, if developing student thinking skills is a central goal, assessments must involve problems and provide opportunities where the complexity of the student's thinking process is exhibited.

In order to discover the correlations between the different variables in the experiment, an analysis of variance (ANOVA) was performed (14). This method identifies the factors that are statistically correlated. The variables studied were whether the students self-selected into a particular section, the section of Chem. 109, the section of Chem. 110, the student octile in Chem. 109, the student gender, the relative rank given by the assessor, the grade in Chem. 110, the time spent out of class, and the time spent in groups. Table 3 summarizes the results of log-linear ANOVA analyses of the course data for two-way and three-way correlations. A two-way correlation between two variables means that the values of one variable helped determine the value of the other, while a three-way correlation means that two variables together were correlated with the value of the third. There were no significant correlations between students who self-selected into specific lecture sections and the previous Chem. 109 lecture section, Chem. 109 octile, gender, or assessor relative rankings. There were also no significant effects of gender on the assessor relative rankings. There were significant two-way and three-way correlations between the Chem. 109 rank in class, Chem. 110 grades, and assessor relative rankings. The two-way interactions indicated that for each lecture, students receiving AB or better were proportionately over-represented among students in the top assessor ranking. The three-way interaction indicates that students receiving AB or better in Chem. 110 and were in the upper half of their Chem. 109 class were over-represented among students in the top assessor ranking. Both two-way correlations between the assessor ranking and the Chem. 109 lecture section or the Chem. 110 lecture section were significant. It is interesting that the Chem. 109 lecture section that produced students with the better assessor rankings was taught by a university teaching award winner. The three-way correlation though was not significant. Some two and three-way interactions between time spent, lecture section, and assessor ranking were fairly significant. A two-way interaction indicated that proportionately more students in the SAL section spent over 15 hrs./week out of class than in the RL section. Interestingly, there was also an over-representation of the 0-7 hrs./week students in the top rank. The three-way interaction indicated that among the >15hrs./week students, a student in the RL section was more likely to be ranked second than first compared to a SAL student. Finally, the time spent working with others did not correlate with the assessor rankings.

Effects on Faculty Evaluations

The changes to an active learning format initiated four years ago caused significant changes in the student course evaluation forms of the SAL instructor between the 12 times it was taught with lectures and the 4 times it was taught with SAL. Answers to the questions- "was the course interesting?" and "was the instructor effective?"- improved significantly from 3.87 to 4.36 and 4.22 to 4.55, respectively, on a 5 point scale. There were no significant changes in the student evaluations of the instructor's preparation level, the background assumed for the course, or the pace of the course. The amount of problem work assigned did change significantly from 2.91 to 3.24 on a 5 point scale, where 3 is labeled "about right".

Conclusions

This study was designed to identify whether differences in student competence resulted from using a SAL strategy. The definition and measurement of student competence in the two sections are controlled by individual faculty assessors in client departments, so the assessment embodies the range of competence definitions that one would expect in a large university. The forced relative assessor rankings hid the fact that the assessors felt almost all the students were quite competent, and it accentuated the differences that resulted from the SAL experiences, especially if the faculty used competence criteria that monitored the sophistication of the student problem solving strategies. The student survey data indicates the students felt the performance differences were related to their ability to demonstrate what they had learned and their preparation levels. The faculty assessor data indicates the differences were related to the thinking skills the students exhibited.

It is important to recognize that the goals for the two lecture sections were different, and that this study was designed to test the attainment of the SAL section's goals, i.e., improved student competence in thinking and solving new problems. Several caveats are in order. The oral exam format did not stress the specific material actually covered in Chemistry 110, in particular the very detailed, graphically oriented and spreadsheet based problem solving method that was emphasized in the RL section. The 25 faculty were by design not made aware in advance of the specific differences in method or content between the two sections. There was no common written examination to test relative performance on the specific course material. There is no reason to believe that the SAL students would do better or worse on traditional examinations. In fact, when the SAL instructor gave traditional timed examinations with challenging quantitative problems from old examinations questions when his classes did not use cooperative methods, the student performance was not significantly different. There are many other questions that the study did not answer: are the differences sustained, were the effects instructor dependent, can others repeat the results, is the content knowledge improved as well? These questions must await further work.

This study has important implications for efforts to change learning processes and experiences, because it shows that one can achieve demonstrable improvements in student performance as judged by faculty colleagues. It also shows that assessment strategies in structured active learning settings need to provide opportunities for students to exhibit their thinking skills. A separate consequence of the assessment was that 25 faculty from across the university became more interested in active learning methods as a result of interacting with enthusiastic students. It is important for other faculty to try similar experiments, using their insights and ideas in order to discover the methods that will optimize science education.

Acknowledgments: This work was sponsored by the National Science Foundation under grant DUE-9455928, the Advanced Research Projects Agency under grant EEC-8721545, and the College of Letters and Science at the University of Wisconsin. The authors would like to express their gratitude to the students in the two courses and the 25 faculty volunteers whose generous donation of time made this study possible.

REFERENCES

1. Rubinstein, E., *Science* **1994,** *266*, 843-875 .

2. Barr, R. B.; Tagg, J., *Change* 1995, *November/December 1995*, 13-25 .

3. Tobias, S., Ed., *They're Not Dumb, They're Different* **1990**, 1st ed. (Research Corporation, Tucson).

4. Tobias, S., Ed., *Revitalizing Undergraduate Science* **1992,** 1st ed. (Research Corporation, Tucson).

4. Newmann, F. M., *Phi Delta Kappan* **1991,** *72*, 458-463 .

5. MacGregor, J., in *New Directions for Teaching and Learning* **1990**, *42*, pp. 19-30. (Jossey-Bass Inc..)

6. Schoenfeld, A. H., in *Cognitive Science and Mathematics Education*, , **1987**, pp. 189-215, 1st ed., A. H. Schoenfeld, Ed. (Lawrence Erlbaum Associates, Hillsdale, NJ).

7. Heller, P.; Keith, R.; Anderson, S. , *Am. J. Phys.* **1992**, *60*, 627-636.

8. Heller,P.; Hollabaugh, M., *Am. J. Phys.* . **1992**, *60*, 637-644.

9. Wright, J. C., *J. Chem. Educ* **1996**, *73*, 827-832.

10. The faculty assessors and departments were Biochemistry- A. Attie, T. Martin, D. Nelson; Biomolecular Chemistry- P. Bertics; Chemical Engineering- K. Bray, D. Cameron, J. DePablo, C. Hill, S. Kim, T. Root; Clinical Chemistry- M. Evenson; Geology- L. Baumgartner, J. Valley; Materials Science- S. Babock, E. Hellstrom, R. Matyi; Mathematics- S. Baumann, M. Certain, T. Millar, A. Nagel; Pharmacy- R. Burnette, K. Conners, C. Royer, G. Zografi; Soil Science- P. Helmke.

11. Patton, M. Q. *Qualitative Evaluation and Research Methods*, **1990,** 2nd ed. (Sage,

Newbury Park, CA).

12. A. Coffey and P. Atkinson. *Making Sense of Qualitative Data: Complementary Research Strategies* **1996** (Sage Publishing, Thousand Oaks, CA).

13. Harris, D. C., *Quantitative Chemical Analysis* **1995**, 4th edition (Freeman).

14. Johnson, R., *Elementary Statistics* **1988** (PWS-Kent Publishing Co., Boston, MA).

TABLE 1

Faculty and Student Questionnaire Results

	Mean SAL	Mean RL	p Value
Faculty Questionnaire (1 disagree-6 agree)			
1) Student felt at ease	4.74	4.70	0.89
2) Student is well-prepared for other science courses	4.84	4.48	0.045
3) Confidence that student performance on oral reflects true competence	4.68	4.45	0.22
4) This student demonstrated overall competence	4.79	4.17	0.0013
Relative Rank (1 is most competent)	3.68	4.80	0.0066[1] (0.023)[2]
Adjusted Absolute Rank[3] (1 is most competent)	1.77	2.22	0.0002
Student Questionnaire (1 disagree-6 agree)			
1) I felt at ease	4.90	4.95	0.64
2) I demonstrated what I learned	4.63	4.18	0.026
3) Demonstrated ability to relate knowledge in new contexts	4.68	4.36	0.12
4) I was fluent in responding to questions	4.39	3.91	0.014
5) Demonstrated I am knowledgeable in chemistry	4.74	4.27	0.0066
6) I appeared nervous	3.06	2.93	0.54

7) I feel well-prepared for other science courses	4.99	4.42	0.0005
8) Compared to other college courses, Chem.110 is at the top	4.77	4.05	0.0027
9) What grade would you assign yourself based on the competence you demonstrated?[4]	88.5	78.6	<0.001
10) How many hours did you spend out-of-class in Chem. 110?[5]	3.96	3.47	0.024
11) What portion of time did you spend working with other students on out-of-class work[6]	3.43	2.32	<0.001

[1] p value based on Wilcoxon Matched Pair Sign Rank Test.

[2] p value based on Signed Test.

[3] Faculty placed students on an absolute scale, similar student performances were grouped together, and the groups were ordered in rank from first to last.

[4] Grade based on 0-100 scale.

[5] Answers were 1) 0-3; 2) 4-7; 3) 8-11; 4) 12-14; 5) 15-17; 6) 18-20; or 7) more than 20 hours/week.

[6] Answers were 1) 0-20; 2) 20-40; 3) 40-60; 4) 60-80; or 5) 80-100%

TABLE 2

Difference in Student Performance for Different Assessor Approaches to the Oral Discussions.

Category Sub-category	**# of Assessor Faculty**	**Mean relative rank-SAL**	**Mean relative rank- RL**	**p-Value- Matched pair sign rank test**
OUTCOMES	10	4.2	4.5	>0.37
Analogy	6	4.4	4.7	>0.60
Analysis	4	3.5	4.2	>0.28
PROCESS	15	3.3	5.0	<0.01
Meta-awareness	11	3.3	5.3	<0.01
Agility	4	3.2	4.2	>0.27

TABLE 3

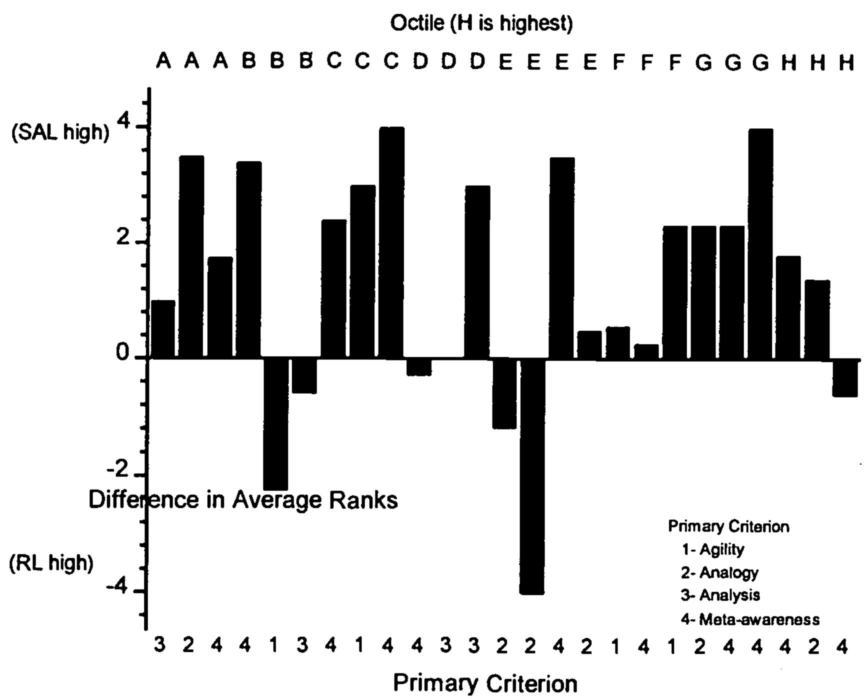
Results of Analysis of Variance and Correlation between Experimental Variables

FACTOR 1	**FACTOR 2**	**FACTOR 3**	**p VALUE**
1. SELF-SELECTION EFFECTS			
self-selection	109 lecture		>0.31
self-selection	octile		>0.23
self-selection	gender		>0.38
self-selection	110 lecture	assessor rank	>0.81 for RL

			>0.17 for SAL
2. GENDER EFFECTS			
gender	assessor rank		>0.85
3. GRADE EFFECTS			
109 rank in class	110 grade		<0.05
109 rank in class	assessor rank		<0.05
110 grade	assessor rank		<0.05
109 rank in class	110 grade	assessor rank	<0.05
4. CHEM. 109 EFFECT			
109 lecture	assessor rank		<0.05
110 lecture	assessor rank		<0.05
5. STUDENT EFFECTS			
time spent out of class	assessor rank		<0.10
time spent out of class	110 lecture		<0.10
time spent out of class	110 lecture	assessor rank	<0.10
time spent in groups	110 lecture	assessor rank	>0.64

FIGURE 1

The bars show the difference in student ranks for each assessor between the SAL and the RL students. The letters above the bars indicate the octile of students that a given assessor interviewed with A being the lowest and H the highest octile. The numbers below the bars indicate the classification of the criterion each assessor used according to the code indicated in the figure.



NOTICE

REPRODUCTION BASIS

☑ This document is covered by a signed "Reproduction Release (Blanket)" form (on file within the ERIC system), encompassing all or classes of documents from its source organization and, therefore, does not require a "Specific Document" Release form.

☐ This document is Federally-funded, or carries its own permission to reproduce, or is otherwise in the public domain and, therefore, may be reproduced by ERIC without a signed Reproduction Release form (either "Specific Document" or "Blanket").