The Effectiveness of Using Calibrated Peer Review (CPR) Against Non-CPR (Traditional) Means in Submitting Chemistry Laboratory Reports

W.N. William* and J.K. Ho

*Department of Chemistry and Chemical Biology University of New Mexico, Albuquerque, NM 87131, USA. *Current address SEKOMU P.O. Box 370, LUSHOTO TANZANIA: Tel: +255716457062 E-mail address: wwillia1@unm.edu or william.wilson@sekomu.ac.tz*

Abstract: *This paper examines the impact of the use of CPR in submitting general Chemistry (123L) laboratory report. This is expected to improve writing skills and alleviate grading burdens particularly when dealing with a large class due to lack sufficient instructors and high grading burden. Analysis of Variance (ANOVA) and t-test were used in statistical analyses. When ANOVA was used for group I students (11 sections) post-laboratory reports submitted by using CPR revealed $F = 0.87$, $p > 0.01$, which implies it is consistent with the null hypothesis. The ANOVA done on group II (15 sections) post-laboratory using CPR revealed $F = 2.07$, $p > 0.01$, which is statistically significant. On the other hand, the comparison of students' who did post-test after using CPR and Non-CPR user revealed $t = 4.18$, $p < 2 \times 10^{-5}$, $t = 6.3$, $p < 7 \times 10^{-10}$, which are statistically significant respectively. In addition, comparison using ANOVA for group I who did pre-test and post-test after using CPR and group II which did not use CPR revealed $F = 2.94$, $p < 3 \times 10^{-5}$, $F = 2.20$, $p < 4 \times 10^{-4}$, which are statistically significant respectively. It is most probable that the noted achievements may not necessarily be due to the use of CPR because the time spent in this research and size of sample used. Indeed, both t-test, and ANOVA analyses have shown existence differences between pre-test and post-test scores, regardless of whether or not the group used CPR to submit post-laboratory report. Statistical analysis has provided little support to connect the use of CPR programme and student writing skill improvement.*

INTRODUCTION

In this rapidly changing world, one of the most important challenges in education is the development of effective techniques that help students learn difficult concepts through writing. Since the College Board offered its first writing examinations in 1901, there have been obvious declines in the writing abilities among students at different educational levels (Breland and Gaynor, 1979). Over the decades, the effort to improve student writing skills has been a principal focus of secondary and college-level education. The literature shows that regardless of the type of discipline, students gain a more profound understanding of conceptual material when they write about what they are learning (Zinsser, 1988).

In general, the teaching of writing can enable instructors to meet the essential teaching challenges that include (1) assisting students in developing conceptual understanding of subject matter (2) enhancing communication skills, and (3) providing opportunities for critical thinking (Kovac and Sherwood, 1999).

Indeed, teaching writing to students with diverse educational backgrounds, abilities and interests is a difficult task for faculties in larger Colleges and Universities, especially in introductory science courses. Many of these challenges are not adequately especially in introductory science courses. Many of these challenges are not adequate addressed through lectures, discussions, laboratories, and problem sets that are part of the traditional first-year college-level science course.

Teaching of writing techniques has been traditionally used in chemistry laboratory course to reinforce student engagement in concepts and critical thinking (Kovac and Sherwood, 1999; Kovac and Sherwood, 2001; Rosenthal, 1987; Oliver-Hoyo, 2003; Schepmann and Hughes, 2006). Furthermore, it has been observed that the writing of technical reports is often poor and students lack motivation to improve their writing (Doody and Gibbens, 1954).

There are undeniable challenges specifically related to *large classes* using the laboratory in submitting laboratory reports, which include high grading burdens and lack of sufficient instructor expertise in creating and evaluating writing assignments (Hobson and Schafermeyer, 2006; Walvoord, et al., 1997). The idea of ñpeer reviewò has been employed as a new approach for teaching oral and written communication skills, and is grounded in the philosophies of active learning (Pieget, 1971). Peer review reduces an instructorôs grading load and maximizes the use of teaching resources (Bound, 1988; Falchikov, 1986).

Through peer review, students build their writing skills from the joint construction of knowledge through writing reviews of each otherôs work. In freshman Chemistry laboratory courses, students from diverse backgrounds must gain skills in scientific writing and critical thinking at the earliest stage of their college careers which may be the best suited for the application of peer-review teaching methodologies. The faculty of science (Chemistry or Biology) interested in implementing writing in a laboratory course usually face many challenges (Kovac, 1999), which include (1) design of effective assignments, (2) knowing how and where to use writing within the laboratory curriculum, and (3) how to provide feedback that improves the studentsô understanding of content and critical thinking.

Non-CPR (traditional) method is at the moment the most commonly easy to used in both teaching and writing essay reports by the university students. In the eyes of the user this can assess all levels of learning objectives. It provides the chance for an individual to show originality and creativity. Also writing report in essay format furnishes students with valuable writing practices (Chan, 2009 and Watson, 2001). Some disadvantages of non-CPR essay report writing include longer time to provide students feedback particularly for large class. A maker (lecturer) can have some biases in grading influenced by various variables (gender, handwriting, length of response etc). Also due to subjectivity nature of essay writing reports, grading is unreliable even for the same assessor at different period of time. In addition, students spend a lot of time in writing essay type of reports (Chan, 2009 and Watson, 2001).

Over the past two decades, a number of chemistry educators have demonstrated the value of a specific peer review technique referred to as ñcalibrated peer reviewò (CPR). CPR is a web-based delivery tool that manages the submission and evaluation of students' written assignments (Russell, 2005). It was developed in 1998 to foster science literacy, constructivist learning and critical thinking in introductory chemistry classrooms in California Colleges and universities (Russell, et al., 1998). CPR is a comprehensive approach which not only promotes student understanding through frequent writing assignments, but also develops student critical thinking skills through the process of peer-review and self-review. Furthermore, it has been noted that CPR improves students writing skills in all disciplines including chemistry lecture assignments in many countries (Schepmann, 2006; Shibley, et al., 2001; Nilson, 2003; Guilford, 2001) and laboratory reports (Kovac and Sherwood, 2001; Rosenthal, 1987; Hollenbeck, et al., 2006; Fawkes and Berry, 2001; Widstrand, et al., 2001; Koprowiski, 1997; Barnett and Blumner, 1999). CPR enables students to develop a number of skills that include: abstracting, persuading, developing logical arguments, describing, assessing, criticizing, analyzing, and reviewing (Hollenback, et al., 2006). These are important higher-order thinking skills as described by Bloom (Bloom, 1956). CPR is useful in a wide range of disciplines, varying from small (20) and large (500) number of students and levels of education (Russell, 2005).

Based on previous studies, students from Chemistry, Biology, Economics, and Physiology courses taught using CPR have performed approximately 10% better on the examinations than students taught through traditional methods (Russell, 2001; Chapman, 2000). This finding was the same for all students who participated in the studies, and was independent of the type of examination questions given in the courses (Hollenbeck, 2006, and Russell, 2005).).

The use of the CPR programme is growing rapidly. For example, in the year 2001, there were 101 universities and colleges that used CPR. During this time, CPR served more than 520 courses, enrolled more than 16,000 students, and had about 175 library assignments in its data base (Murphy, 2001). In the year 2004, the CPR user report showed that 500 institutions used CPR to support 1900 courses, which had a total enrollment of over 72,000 students (Murphy, 2001). In 2005, the report showed that CPR has been adopted by over 800 institutions and served more than 120,000 students (Hollenbeck, 2006). More importantly, the assignment library of CPR has expanded exponentially to 1275 assignments.

Quantitative studies have revealed that students who learn by using CPR programme have improved test scores over those who are taught using traditional lecture approach (Russell, 2001; Chapman, 2000; Palaez, 2002). Other researchers suggest that CPR led to improvement of studentsô performance in both essay writing and critical thinking (Heise, 2002; McCarty, et al., 2005; Carlson, et al., 2003; Carlson and Berry, 2005; Donovan, 2003; Robinson, 2001). Some supportive literature suggested that well-written CPR assignments can facilitate course content

mastery (Furman, 2003; Gerdeman, et al., 2007; Kim, 2005; Margerum, et al., 2007), and can be as effective as other methods of teaching writing skills (Heise, 2002; Plutsky and Wilson, 2004; Palaez, 2001).
Conversely, other studies have disputed the value of CPR. Walvoord, et al., 2008 reported that CPR did not improve studentsô technical writing skills, nor did it improve their scientific understanding of written summaries from publications (Walvoord, et al., 2008). An analysis carried out by Reynolds, et al., 2008 revealed that the use of CPR did not match the expectations to improve student learning. Because of such conflicting views we fell it is necessary for us to carry out our own research to determine the effectiveness of using CPR in the context of our own course (CHEM 123 L) for improving writing skills.

Methods

This study was designed to determine the effectiveness of implementing the CPR programme in teaching of a first-semester, general chemistry laboratory (CHEM 123L). The course contained a total of eight experiments with 26 sections of students. Eleven sections (Group I) carried out the first four experiments using the CPR programme and then reverted to the traditional approach for the remaining four experiments. Fifteen sections (Group II) carried out the first four experiments with traditional approach and then used CPR programme for the remaining four experiments. Students did a pre-test before doing each experiment afterwards they conducted the experiment. For those experiments conducted by a group which did not use CPR, the students submitted their post-lab reports through the traditional hard-copy approach for grading. For the group which used CPR, the students submitted their post-lab reports through the CPR programme afterwards the group was given a post-test. The performances on the pre-test and the post-test were then compared within each group.

At the beginning of the course, the Teaching Assistants (TA) or instructor gave a presentation to the students on the general concepts of CPR and how it works. The presentation included information about instructorœs ability to see all comments and reviews made by the students. The laboratory experiment was done in three hours. Students who submitted a traditional post-lab report handed in their report within seven days. Students who submitted their post-lab report through the CPR programme had approximately two weeks to complete the assignment. All essay questions were instructor-developed items for topics addressed in the experiment. Each assignment was designed with goals, source materials, guiding questions, and three example essays to help students ñcalibrateò their ability for peer review. The CPR system automatically managed the peer review process.

Data Collection

The grading system for the CPR process was assigned a total of 20 Marks with following breakdown: text entry submitted through the CPR programme consists of the following break down: 70% of the grade was given for the quality of report as entered by the student; 13% of the grade was based on the reviewing process in the calibration performance; 12% of the grade was based on the review of the studentœs classmate's work and 5% of the grade was based on self-evaluation. The remaining

80 marks of the course grade were based on TA or instructor evaluations on general performance of both pre-lab and post-lab questions reports.

To assess the value of CPR, the following question was asked at the beginning as pre-test of the course and after every completion of the four experiments using CPR programme the same question was used as post-test. The grading rubric was used to award 10 marks.

The following is a studentøs laboratory procedure, observation and experimental data from the preparation of a 100 mL of 0.4 M glucose solution. Comment on this studentøs preparation of the solution. Describe the problems of preparation and how to correct the problems. Write the answers in essay format with clear logic and correct spelling and grammar (10 Marks).

Making 100 mL of a 0.4 M sugar solution	
Experiments	**Observations**
1. Prepare a clean and dry, 100 mL volumetric flask with a stopper.	1. Obtained a volumetric flask from the drawer. Cleaned it with soap and dried it with paper towel.
2. Weigh 720 g of glucose by using a weigh boat and record the mass. 0.4 M x 100 mL = 40 moles 40 moles x 180 g/mole = 720 g	2. The mass of glucose is 719.980 g = (740.01 g ï 20.03 g of weigh boat).
3. Transfer all the glucose into volumetric flask	3. Although I was very careful, I lost some glucose during the transfer. It was hard not to spill it; the flask opening is too small.
4. Add the water to the volumetric flask to the mark.	4. The water line was just below the mark.
5. Shake the flask to make all sugar dissolve.	5. A little water was leaking out of the stopper during the mixing. It is very hard to dissolve the entire solid. It took me almost 40 minutes to dissolve it.

Data Analysis for Post-lab (Reports Submitted Using CPR) and Pre-est/Post-test Scores

The scores for all eight post-lab reports obtained through the CPR process and the pre-test and post-test scores were analyzed using ñStudentøsò t-test, and Analysis of Variance (ANOVA) (Devore and Farnum, 2005). The t-test was used to confirm whether the mean of two normally distributed scores are equal. Our null hypothesis states that the mean values from pre-test and post-test are not statistically different (i.e., $H_o : \mu_1 = \mu_2$).

The ANOVA-single factor analysis was used in comparing the mean of more than two samples to determine whether differences in the means are real or just random errors. This was done in a single test with ANOVA rather than pair wise comparisons using the t-test approach. ANOVA helped to compare the means and standard deviations of more than two groups (in these case scores on pre-test and post-test for students who used or did not use CPR) and determined whether there was significance difference in variance from one group to another. Our *null hypothesis* for ANOVA was that all pre-test and post-test variances are equal (i.e., $H_o : \sigma_1^2 = \sigma_2^2 = \sigma_3^2 = ... = \sigma_k^2$) and k stands for a set of scores in a give section.

Results and Discussion

Comparison of CPR Post-lab Reports Submitted by Groups I and II

First, we compared the scores of the post-lab reports graded by the CPR program for all sections in Group I. ANOVA was used to analyze average score of the best post-lab reports (Group I) out of four submitted reports using CPR. This analysis revealed that variance in scores between the individual sections in Group I was not significant (see **Table 1**), indicating that the overall performance of this group was approximately the same (F = 0.87, p > 0.57 at 0.05 alpha level), which is consistent with the null hypothesis ($H_o : \sigma_1^2 = \sigma_2^2 = \sigma_3^2 = ... = \sigma_k^2$).

Table 1. One way Analysis of Variance of Average Post-lab Scores for Group I Students Who Submitted Assignments 1-4 Using the CPR Programme

Sections	No. of Students	Sum total of scores	Average scores	Variance
12	19.00	303.02	15.95	11.92
15	15.00	242.08	16.07	12.22
16	14.00	232.09	16.58	6.35
20	21.00	340.98	16.24	14.54
24	17.00	291.61	17.15	2.82
27	21.00	344.00	16.38	9.12
28	11.00	170.60	15.51	16.53
29	20.00	345.34	17.27	3.04
30	20.00	351.80	17.59	3.17
31	19.00	298.37	15.70	12.83
34	16.00	259.85	16.24	7.60

ANOVA

Source of Variation	SS	df	MS	F	P-value	F-crit.
Between Groups	76.90	10.00	7.69	0.87	0.57	1.88
Within Groups	1614.70	182.00	8.87			
Total	**1691.61**	**192.00**				

Comparison of Group II scores variances (see **Table 2)** revealed the existence of a significant difference between the sections in this group F-value being significant at (F = 2.07, p < 0.01, 0.05 alpha level), and were consistent with the rejection of the null hypothesis ($H_o : \sigma_1^2 = \sigma_2^2 = \sigma_3^2 = ... = \sigma_k^2$). This result could have been due to a variation in student ability or variation in the quality of the TAs as well as instructor. However, when sections 18 and 37 were excluded, the variances in scores decreased significantly. Without these two sections, $H_o : \sigma_1^{\ 2} = \sigma_2^{\ 2} = \sigma_3^{\ 2} = \ldots\ldots = \sigma_k^{\ 2}$ for Group II the null hypothesis is accepted.

Table 2: One-way Analysis of Variance of Average Post-lab for Group II Scores of Students Who Submitted Assignments 5-8 Using the CPR Programme

Sections	No. of Students	Sum total of scores	Average scores	Variance
11	21.00	355.40	16.92	5.11
14	15.00	240.06	16.00	4.62
17	18.00	298.28	16.57	4.43
18	17.00	228.74	13.46	20.59
21	17.00	277.74	16.34	7.67
22	20.00	319.89	15.99	13.27
23	19.00	312.94	16.47	12.57
25	18.00	311.20	17.29	2.91
26	20.00	318.28	15.91	10.17
32	18.00	273.17	15.18	8.03
33	22.00	337.46	15.34	9.98
37	15.00	205.37	13.69	17.44
38	15.00	223.00	14.87	10.70
39	20.00	306.35	15.32	9.22
41	11.00	173.14	15.74	9.89

ANOVA

Sources of variation	F	p-value	F-crit.	Comments
All sections	2.07	0.01	1.73	There is significant difference
Excluding section 37	1.79	0.05	1.76	There is significant difference
Excluding section 18	1.61	0.08	1.76	There is insignificant difference
Excluding sections 18 and 37	1.10	0.36	1.80	There is insignificant difference

For the purpose of comparison of individual student pre-test and post-test scores, we carried out different analyses on the results using t-test, and ANOVA-single factor. The results are as indicated below:

(i) The t-test

The t-test treatment of an objective essay results was carried out under null hypothesis, which states that both pre-test and post-test compared have an equal mean (i.e., $H_o : \mu_1 = \mu_2$). Furthermore, it is conceivable to assume that post-test scores would be expected to be greater than pre-test scores. The results in **Table 3** and **Table 4** are based on t-one tail that is directional. Through comparison of pre-test and post-test scores for Group I students who submitted assignment 1-4 using CPR, the results from the t-test are summarized in **Table 3**. The table consists of four out of eleven sections (15, 16, 20, and 24), which were shown to have significant differences between pre-test and post-test score means. These four sections have mean values that are above the theoretical t-values as shown in the histogram in **Figure 1**. Nevertheless, the summation of all sections scores (pre-test and post-test for students who used CPR) revealed a t-test value ($t = 4.18$, $p < 2 \times 10^{-5}$) that was statistically significant. Thus, the null hypothesis $H_o : \mu_1 = \mu_2$ is rejected.

Table 3: The t-test Results (Pre-test-post-test Comparison) for Group I Students who did the Post-test After Using CPR

Section	Students	$\overline{X_1}$	$\overline{X_2}$	t-test calculated	t-one tail	t-two tail	P one tail	P two tail
12	16	2.25	3.28	1.50	1.69	2.04	0.07	0.14
15	11	2.82	4.82	2.37	1.72	2.08	0.01	0.02
16	8	3.13	4.69	2.14	1.76	2.14	0.02	0.05
20	19	2.89	4.18	2.47	1.69	2.03	0.01	0.02
24	13	3.50	4.58	1.73	1.71	2.06	0.05	0.10
27	18	4.06	4.61	0.90	1.69	2.03	0.19	0.37
28	15	4.53	4.00	0.69	1.70	2.05	0.25	0.50
29	22	4.52	4.75	0.40	1.68	2.02	0.35	0.70
30	17	3.68	4.32	1.43	1.69	2.04	0.08	0.16
31	16	3.66	3.75	0.21	1.70	2.04	0.42	0.83
34	11	2.86	2.68	0.29	1.72	2.09	0.39	0.78

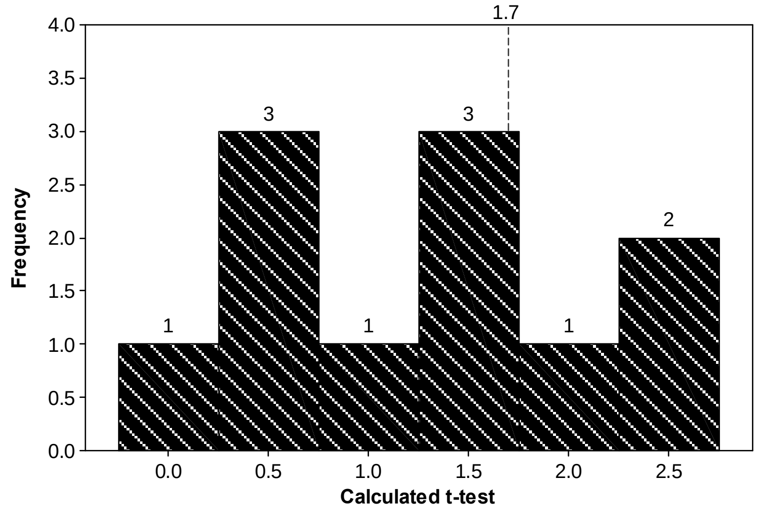

Figure 1. The t-test Statistics Frequency Histogram for Students Performance on Pre-test-post-test after Submission Assignments 1-4 Using CPR with Theoretical Value of t at 1.7 (t-one tail)

Table 4: The t-test Results (Pre-test-post-test Comparison) for Group II Students who did Post-test Without the Use of CPR

Section	Students	$\overline{X}_1$	$\overline{X}_2$	t-test calculated	t-one tail	t-two tail	P one tail	P two tail
11	20	3.28	3.95	1.40	1.69	2.02	0.08	0.17
14	15	2.90	4.23	1.99	1.70	2.05	0.03	0.06
17	16	3.22	4.16	2.06	1.69	2.04	0.02	0.05
18	18	2.33	3.31	1.87	1.69	2.03	0.03	0.06
21	15	2.90	3.87	1.42	1.70	2.05	0.08	0.16
22	14	3.29	4.32	2.37	1.70	2.05	0.01	0.02
23	16	3.50	3.06	0.90	1.69	2.04	0.18	0.37
25	17	3.65	4.38	1.54	1.69	2.03	0.06	0.13
26	20	3.50	2.50	0.05	1.68	2.04	0.48	0.96
32	18	2.72	3.42	1.77	1.69	2.03	0.04	0.08
33	15	2.90	3.93	1.82	1.70	2.05	0.04	0.08
37	15	2.93	4.27	2.04	1.70	2.05	0.02	0.05
38	13	2.54	3.85	1.90	1.71	2.06	0.03	0.07
38	22	4.41	4.11	0.49	1.68	2.01	0.31	0.62
41	11	3.23	4.41	1.72	1.72	2.08	0.12	0.24

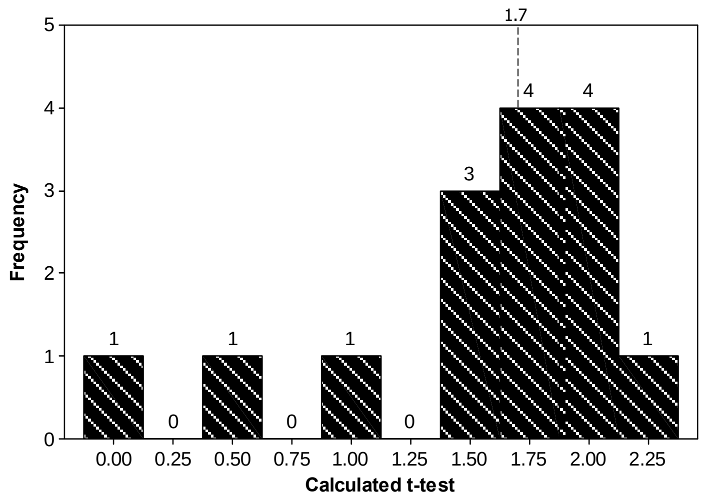

Figure 2: **The t-test Statistics Frequency Histogram for Student Performance on Pre-test-post-test without Use of CPR Theoretical Value at 1.7 (t-one tail)**

Among students who did not use CPR, the t-test revealed significant differences between pretest and posttest scores from nine out of fifteen sections (**Table 4** and **Figure 2**). On other hand, the t-test analysis performance on the sum of all sections scores (pre-test and pos-test for students who did not use CPR) yielded a value ($t = 6.30$ $p < 7 \times 10^{-10}$) that was statistically significant. Thus, the null hypothesis $H_o : \mu_1 = \mu_2$ is rejected.

(ii) ANOVA

ANOVA-single factor analysis was performed on the Group I and Group II pre-test and post-test) scores. This test was applied based on the null hypothesis that all scores in each section have equal variance (i.e., $H_o : \sigma_1^2 = \sigma_2^2 = \sigma_3^2 = ... = \sigma_k^2$). A separate analysis was performed to identify difference in performance between Group I and Group II. The calculated F-values indicate the existence of significant differences in variances for the two groups, which can be seen in **Table 5**, ($F = 2.94$, $p < 3 \times 10^{-5}$) and **Table 6** ($F = 2.20$, $p < 4 \times 10^{-4}$). Thus, the null hypothesis $H_o : \sigma_1^2 = \sigma_2^2 = \sigma_3^2 = ... = \sigma_k^2$ is rejected for both Group I and Group II.

Table 5: One-way Analysis of Variance of Scores for Group I Students Who did Pre-test and Post-test After Using CPR

Groups	Count	Sum	Average	Variance
BF12	16.00	36.00	2.25	3.40
AF CPR	16.00	52.00	3.28	4.17
BF 15	11.00	31.00	2.82	3.76
AF CPR	11.00	51.00	4.68	3.06
BF 16	8.00	25.00	3.13	1.91
AF CPR	8.00	37.00	4.69	2.35

BF 20	19.00	55.00	2.89	1.24
AF CPR	19.00	79.00	4.18	3.92
BF 24	13.00	45.50	3.50	0.92
AF CPR	13.00	59.00	4.58	4.12
BF 27	18.00	73.00	4.06	4.64
AF CPR	18.00	83.00	4.61	3.19
BF 28	15.00	68.00	4.53	3.55
AF CPR	15.00	60.00	4.00	3.46
BF 29	22.00	99.50	4.52	3.08
AF CPR	22.00	104.50	4.75	4.21
BF 30	17.00	62.50	3.68	0.84
AF CPR	17.00	73.50	4.32	2.62
BF 31	16.00	58.50	3.66	0.69
AF CPR	16.00	60.00	3.75	2.40
BF 34	11.00	31.50	2.86	3.35
AF CPR	11.00	29.50	2.68	0.96

ANOVA

Source of Variation	**SS**	**Df**	**MS**	**F**	**p-value**	**F-crit.**
Between Groups	180.29	21.00	8.59	2.94	0.00003	1.59
Within Groups	905.47	310.00	2.92			
Total	1085.76	331.00				

BF: prê-test, AF-CPR : post-test for student who submitted assignment 1-4 using CPR programme

Table 6: One-way Analysis of Variance for Scores of Group II Students who did Pre-test and Post-test Without the Use of CPR

	Count	Sum	Average	Variance
BF 11	20.00	65.50	3.28	1.51
AF-No CPR	20.00	79.00	3.95	3.13
BF 14	15.00	43.50	2.90	2.04
AF-No CPR	15.00	63.50	4.23	4.67
BF 17	16.00	51.50	3.22	1.20
AF-No CPR	16.00	66.50	4.16	2.12
BF 18	18.00	59.50	3.31	2.83
AF-No CPR	18.00	43.50	2.90	2.04
BF 21	15.00	43.50	2.90	2.04
AF-No CPR	15.00	58.00	3.87	4.87
BF 22	14.00	46.00	3.29	0.72
AF-No CPR	14.00	60.50	4.32	1.95
BF 23	16.00	56.00	4.32	1.95

AF-No CPR	16.00	49.00	3.50	2.10
BF 25 AF-No CPR	17.00 17.00	62.00 74.50	3.65 4.38	1.40 2.45
BF 26 AF-No CPR	20.00 20.00	70.50 71.00	3.53 3.55	1.46 3.63
BF 32 AF-No CPR	18.00 18.00	49.00 61.50	2.72 3.42	1.39 2.54
BF 33 AF-No CPR	15.00 15.00	43.50 59.00	2.90 3.93	2.54 2.25
BF 37 AF-No CPR	15.00 15.00	44.00 64.00	2.93 4.27	3.96 2.42
BF 38 AF-No CPR	13.00 13.00	33.00 50.00	2.54 3.85	1.60 4.64
BF 39 AF-No CPR	22.00 22.00	97.00 90.50	4.41 4.11	4.47 3.36
BF 41 AF-No CPR	11.00 11.00	35.50 48.50	3.23 4.41	4.97 5.54

ANOVA

Source of Variation	SS	Df	MS	F	p-value	F crit.
Between Groups	167.88	29.00	5.79	2.20	0.0004	1.49
Within Groups	1207.83	460.00	2.63			
Total	**1375.72**	**489.00**				

BF: pre-test, AF-No CPR: post-test for students who did not use the CPR programme

CONCLUSIONS

Statistical analyses (t-test and ANOVA) of pre-test and post-test scores show significant differences for both CPR-users and non-CPR-users. This implies that the improvement noted in posttest performance is not necessarily due to the use of CPR. Also it is most probable that the variation may be due a combination of studentsô ability or variation of the quality of TA, as well as instructors.

Both t-test and ANOVA analyses have therefore revealed existing differences between pre-test and post-test scores, regardless of whether or not the group used CPR to submit their post-lab reports. The statistical analysis gave us little reason to link the use of the CPR programme and student improvement. The establishment of a comprehensive link would require continuing, long-term research on both students who are using CPR and those who non-CPR users, but due to limited time of study on two groups has shown almost the same results. It is conceivable that long time for more research is still needed on both CPR and non-CPR users before clear distinction can be established to distinguish the two approaches. Indeed, this has some limitations of having the same students in every academic year.

ACKNOWLEDGMENTS

We would like to acknowledge the students in general Chemistry course class during the fall 2010 semester and all teaching assistants. The project was generously supported by the University of New Mexico Department of Chemistry and Chemical Biology USA.

REFERENCES

Barnett, R. W. and Blumner, J. S. (eds). (1999). *Writing Centers and Writing Across the Curiculum Programs*, West Port, CT: Greenwood Press.

Bloom, B. S. (Ed). (1956). *Taxonomy of Educational Objectives, Cognitive Domain Longmans*, Green and Company, New York.

Bound, D. (1988). *Developing Student Autonomy in Learning* 2nd ed. London: Kogan Page.

Breland, H. M. and Gaynor, J. L. (1979). ñA Comparison of Direct and Indirect Assessment of Writing Skillò *Journal of Educational Measurement*, Vol. 16, 119-128.

Carlson, P. A. and Berry, F. C. (2003). *Calibrated Peer Review and Assessing Learning Outcomes.* Paper Presented at the 33rd ASEE/IEEE Frontiersô in Education Conference, Boulder, Co.

Carlson, P. A.; Berry, F. C. and Voltmer, D. (2005). Incorporating Student Peer-review into an Introduction to Engineering Design Course. Paper Presented at the 35th ASEE/IEEE Frontiers in Education Conference, Indianapolis, IN.

Chan, C. (2009). Assessment: Essay, Assessment Resources@HKU, University of Hong Kong [http:/ar.hku.hk].

Devore, J.; Farnum, N. (2005). *Applied Statistics for Engineers and Scientists*, 2nd ed. Green and Company, New York, 402.

Donovan, J. B. (2003). ñMy Computer, my Writing Coachò *Howard Hughes Medical Institute* (*HHMI*) *Bulletin*, Vol. 16, 7-12.

Doody, T. C. and Gibbens, V. E. (1954), *J. Chem. Educ.*, Vol. 31, 8-18.

Falchikov, N. and Goldfinch, J. (2000). ñStudent Peer Assessment in Higher Education: A Meta-Analysis Comparing Peer and Teacher Marksò *Review of Educational Research*, Vol. 70, 287-299.

Falchikov, N. (1986). ñProduct Comparisons and Process Benefits of Collaborative Self and Peer Group Assessmentsò *Assessment and Evaluation in Higher Education*, Vol. 11, 146-166.

Fawkes, K. L. and Berry, D. E. (2001). *Positive Pedagogy*, Vol. 1, 1496-1510.

Furman, B. and Robinson, W. (2003). *Improving Engineering Report Writing with Calibrated Peer Review*. Paper presented at the 33rd ASEE/IEEE Frontiers in Education Conference, Boulder, Co.

Gerdeman, R, D.; Russell, A. A. and Worden, K. J. (2007), òWeb-based Student Writing and Reviewing in a Large Biology Lecture Courseò *Journal of College Science Teaching*, Vol. 36, 46-50.

Guilford, W. H. (2001). *Adv. Physiol. Educ*., Vol. 25, 167-173.

Heise, E. A.; Palmer-Julson, A. and Su, T. M. (2002). ñCalibrated Peer Review Writing Assignment for Introductory Geology Coursesò *Abstract with Programs (Geological Society of America)*, Vol. 34, 345-353.

Hobson, E. H. and Schafermeyer, K. W. (1994). ñWriting-to-Learn in Large Classesò *American Journal of Pharmaceutical Education*, Vol. 58, 423-427.

Hollenbeck, J. J.; Wixson, N. E.; Geske, G. D.; Dodge, M. W; Tseng, T. A.; Clauss, A. D. and Blackwell, H. E. (2006). *J. Chem. Educ*., Vol. 83, 1835-1843.

Kim, S. H.; Wise, J. and Hillsley, M. (2005). Learning Technical Writing Skills Through Peer Review: Use of Calibrated Peer Review in unit Operation Laboratory. Paper Presented at AIChE Annual Meeting, Conference Proceedings, Vol. 27, 23-31.

Koprowiski, J. L. (1997). *Coll. Sci. Teach.,* Vol. 27, 133-142.

Kovac, J. and Sherwood, D. W. (1999). *J. Chem. Educ*., Vol. 76, 1399-1403.

Kovac, J. and Sherwood, D. W. (2001). *Writing Across the Chemistry Curriculum*: An Instructorôs Handbook; Printice Hall: Upper Saddle River. NJ.

Margerum, L. D.; Gulsud, M.; Manlapez, R.; Rebong, R. and Love, A. (2007). ñApplication of Calibrated Peer Review Writing Assignments to Enhance Experiments with an Environmental Chemistry Focusò *J. Chem. Educ.,*Vol. 84, 292-300.

McCarty, T.; Parkes, M. V.; Anderson, T. T.; Mines, J.; Skipper, B. J. and Grebosky, J. (2005). *Academic Medicine*, Vol. 80, 67-73.

Murphy, P. (2001). ñCollaborative Innovations: New UC-Wide Partnerships are Transforming the Teaching and Learning Experiencesò http://www.uctltc. org/news /2001/12/feature-6.htm.

Nilson, I. B. (2003). *College Teaching*, Vol. 51, 34-42.

Oliver-Hoyo, M. T. (2003). *J. Chem. Educ*., Vol. 80, 889-894.

Palaez, N. (2002). ñProblem-Based Writing with Peer Review Improves Academic Performance in Physiologyò *Advances in Physiology Education*, Vol, 26, 174-186.

Palaez, N. (2001). Calibrated Peer Review in General Education Undergraduate Human Physiology. In Proceedings of Annual International Conference of the Association for the Education of Teachers in Science, Costa Mesa. California, 8.

Piaget, J. (1971). *Science of Education and the Psychology of the Child*. Longman: London.

Plutsky, S. and Wilson, B. A. (2004). ñComparison of Three Methods for Teaching and Evaluating Writing: Aquasi-experimental Studyò. *Delta pi Epsilon*, Vol. 46, 50-53.

Reynolds, J. and Moskovitz, C. (2008). *Journal of College Science Teaching*, Vol, 60, 247-456.

Robinson, R. (2001). ñCalibrated Peer Review an Application to Increase Student Reading and Writing Skillsò *American Biology Teacher*, Vol. 63, 474-480.

Rosenthal, L. C. (1987). *J. Chem. Educ*., Vol. 64, 996-1001.

Russell, A. A.; Chapman, O. L. and Wagner, P. A. (1998). *J. Chem. Educ*. Vol. 75, 578-583.

Russell, A. A. (2005). *Growth Charts on History of Uses of CPR* Los Angeles: UCLA.

Russell, A. A. (2001). *The Evaluation of CPR, Prepared for HP e-Education*; Business Development, UCLA.

Russell, A. A. (2005). ñCalibrated Peer Review ™ A Writing and Critical Thinking Instructional Toolò (Invention and Impact): *Building Excellence in Undergraduate Science Technology, Engineering and Mathematics (STEM) Education*. Washington DC: American Association for the Advancement of Science.

Schepmann, H. G.;and Hughes, L. A. (2006). *J. Chem. Educ*., Vol. 83, 1024-1031.

Shibley, I. A., Jr.; Milakofsky, L. M. and Nicotera, C. L. (2001), *J. Chem. Educ*., Vol. 78, 50-56.

Walvoord, B. E.; Hunt, L. L.; Dowling, H. F. and McMahon, J. D. (1997). *In the Long Run*: *A Study of Faculty in Three Writing- Across the Curriculum Programs*. Urban, IL: National Council of Teachers of English.

Walvoord, M. E.; Hoefnagels, M. H.; Goffin, D. D.; Chumchal, M. M. and Long, D. A. (2008). *Journal of College Science Teaching*, Vol. 66, 147-158.

Widstrand, C. G.; Nordell, K. J. and Ellis, A. B. (2001). *J. Chem. Educ*., Vol. 78, 1044-1053.

Watson, D. M. (2001). *Education and Information Technology*, Vol. 6, 251-266.

Zinsser, W. (1988), *Writing to Learn*; Harper and Row: New York.