

Exploring the Potency of Peer Evaluation to Develop Critical Thinking for Tertiary Academic Writing

¹Nor Shidrah Mat Daud, ²Alison Gilmore and ³Helen Elaine Mayo

¹Academy of Language Studies, Universiti Teknologi MARA Malaysia, Malaysia

²College of Education, University of Otago, New Zealand

³College of Education, University of Canterbury, New Zealand

Abstract: Producing high quality, well-polished academic writing at tertiary level has always been difficult for undergraduate students. They do not only need to have a considerable proficiency in the language but also the ability to think critically about the issue before they can produce an acceptable piece of work. One approach that contributes to the development of students' critical faculties is peer evaluation. Peer evaluation requires students to be critical in assessing other students' work, defend their own work and subsequently make improvements. Such an evaluation activity allows revisions and editing of drafts in a collaborative learning environment. However, there is a dearth of study on the effect of peer evaluation on language students' development of critical thinking skills particularly in academic writing. The objective of this paper is to investigate the contribution of peer evaluation in developing students' critical thinking skills. A quasi-experimental study was conducted with Social Sciences students who were taking English for Academic Writing course at the International Islamic University Malaysia. The students were observed throughout the drafting stage of their term paper for a duration of seven weeks (of a fourteen-week academic programme) and the Cornell Critical Thinking Test Level X (CCTT-X) was administered at the beginning of the second half of the semester and at the end of the semester. The grades earned on the term papers were identified. The results of the CCTT-X pre- and post-test were compared to examine the effect of treatment on the participants' development of critical thinking. Correlation between the students' critical thinking skills and academic writing ability was also investigated. The students' perceptions of the peer evaluation activity were later elicited in interviews and a short written survey. The findings of the study indicate that peer evaluation activities did help in developing critical thinking skills, thus improving their academic writing performance.

Key words: Critical Thinking • Peer Evaluation • Self-Evaluation • Peer Review • Tertiary Academic Writing
• Rubric

INTRODUCTION

Tertiary writing differs from secondary school as it requires tertiary students to produce writing that is more critical and academic in nature. Objectivity and conciseness are the goals of such writing and certain stylistic rules and guidelines need to be adhered to. All these entail tertiary students needing to critically examine and be engaged with the ideas and issues of the discipline within which they write considerably more complex and detailed texts using an appropriate academic voice [1]. This is a challenge

for the majority of undergraduate students whose writing lessons at school have not focused on such skills. Instead their earlier writing skills are developed based on learning experiences that emphasise language accuracy and mechanics. At the undergraduate level, however, irrespective of the medium of instruction, the ability to write well does not only depend on the quality of language used but also how clear and well-thought out the ideas are. This scholarly writing requires considerable critical thinking by the writer which is integral to the production of insightful and thought-provoking writing.

The method for the teaching of critical thinking is relevant to writing classes. One practical exercise is peer evaluation which requires students to be critical in assessing others' work, defend their own and subsequently make improvements. It helps develop students' critical faculties [2]. By assessing each other, "students refine their thinking, consider the quality of their communication and acquire a heightened sense of the reader's understanding and interpretation of the writing" [3] (p. 215). Other advantages include provision of immediate feedback [4], multiple comments from peers aside from teachers [5, 6], developing students' self-evaluation ability [7] and encouraging "...more effective knowledge transfer of evaluation process and standards" [8] (p.152).

However, although peer evaluation is a common teaching technique in tertiary classes, research concerning its adoption at tertiary level is still scarce [9]. The majority of the studies available compared the reliability and validity of self- and peer evaluation with that of teacher evaluation [e.g. 10-13]. A few other studies looked at the potential of self- and peer assessment to support students' motivation and promote learning [e.g. 14-16]. Studies comparing the effect of self- and peer evaluation activities or peer evaluation and peer review, respectively, have not yet been found.

This study sheds some light on the possibility of developing critical thinking skills using peer evaluation guided by a rubric in comparison to self-evaluation using the same rubric and peer review using a checklist particularly among undergraduates who were doing *English for Academic Writing* course.

Research Objectives: The specific aims of the study are to:

- Investigate whether there are any differences in the level of development of critical thinking skills when peer evaluation, self-evaluation and peer review are adopted and
- Examine the impact of the different treatments on academic writing ability.

MATERIALS AND METHODS

A quasi-experimental study was conducted to see whether peer evaluation would help improve students' critical thinking skills. The design used non-equivalent pre-post-test groups.

The Setting and the Participants: The study was conducted at the International Islamic University Malaysia (IIUM), a government university in Malaysia. *English for Academic Writing* course (LE 4000), offered by the Centre for Languages and Pre-University Academic Development (CELPAD), is compulsory for all undergraduate students of the University as a graduation requirement. Critical thinking skills development is heavily stressed in the course objectives and learning outcomes, thus the relevance of the study for such classes.

Four intact groups of undergraduate students doing the *English for Academic Writing* (LE 4000) at the University were involved in the quasi-experimental study. One group was the control group and three groups were the experimental groups. The experimental groups were peer review, self-evaluation and peer evaluation groups. The control and peer review groups were taught by instructor 1, while the self-evaluation and peer evaluation groups were taught by instructor 2.

To deal with selection threat, the students were those in the same year and from the same discipline of study, which were the social sciences. The instructors were selected by considering their gender, educational background and teaching experience. Details of the participants are as in Table 1 below.

The Procedure: After consent was sought from the participants, the instructors were briefed on the study including training them on how to adopt the rubric for peer and self-evaluation and the checklist for peer review. In the classrooms, the students were also trained on how to use the rubric and the checklist to help increase their confidence [17]. The Cornell Critical thinking Test (CCTT-X) was administered at the beginning of the semester.

Later, the three experimental groups underwent different treatments in addition to writing instructions:

- Peer evaluation with the use of a rubric developed by the primary researcher. The six-point rubric labeled by a scale of *emerging*, *developing* and *mastering* consisted of 12 evaluation criteria. A commentary space was also provided for the students to note down their comments on the work being evaluated.
- Peer review guided by a checklist developed based on the rubric. All the 12 evaluation criteria in the rubric made the 12 reviewing points for the students. No grading work was involved. The checklist also includes a commentary space for students' use.
- Self-evaluation with the use of the same rubric as peer evaluation.

Table 1: The distribution of participants

		GROUP				TOTAL (N=99)
		Instructor 1		Instructor 2		
		CG (N=24)	PR (N=24)	SE (N=24)	PE (N=27)	
FACULTY	Economics	20	0	0	18	38
	Human Science	0	0	24	0	24
	Law	0	24	0	0	24
	Accountancy	4	0	0	9	13
NATIONALITY	Malaysians	21	23	21	22	87
	Others*	3	1	3	5	12
GENDER	Male	4	1	4	10	19
	Female	20	23	20	17	80

Note:

CG = Control Group

SE = Self evaluation

PR = Peer Review

PE = Peer evaluation

*International students from Montenegro, Brunei, Singapore, Indonesia, Maldives, India, Kenya and Yemen.

Peer review group was included to allow comparison with peer evaluation group on the peers' critical thinking development without and with the rubric respectively, while self-evaluation group provided evidence as to whether critical thinking skills would accelerate faster when the rubric was used for peer rather than self-evaluation. The control group learned academic writing without intervention. This means that students received writing instructions from the instructor and then produced essays to be checked and graded by their instructor.

The intervention period started during the three drafting stages of the course project which was in the second half of the semester. Prior to the interventions, the students individually selected their own topic to write about for the course term paper. The term paper was to be approximately 2500 words long. The three drafting stages were:

First : Introduction + First argument

Second : Second argument + Counter argument + Refutation

Third : Conclusion + Abstract + Bibliography

Before each draft was submitted to the instructor to check, students in peer evaluation and peer review groups worked in small groups of three or four to read and orally discuss each other's work. The peer evaluation group, however, had an additional task to do that is rating using the given rubric. These two groups were given approximately 20 minutes to check each draft. Students in the self-evaluation group worked individually to assess and rate their own paper within five minutes. The final draft was submitted at the end of the semester. Observation of the activities was done.

Towards the end of the semester, after the project final drafts had been submitted, the CCTT-X was again administered on the students. A survey questionnaire was also distributed to examine the extent to which they perceived the learning activities assisted their learning. Semi-structured interviews were later conducted with five randomly selected students from each group to get a more detailed feedback on the different learning activities.

The Cornell Critical Thinking Test Level X (CCTT – X):

The students' critical thinking level was investigated using the Cornell Critical Thinking Test Level X (CCTT-X) developed by two Cornell University professors, Robert Ennis and Jason Millman in 1985 [19]. It has high reliability estimates with various populations which ranged from 0.87 to 0.91 [18]. CCTT-X consists of 71 multiple-choice items (a 50-minute test) and is meant for 4th through 14th graders but has been used with undergraduates who are less sophisticated [19]. Four skills are tested in this test instrument which are:

- Deduction
- Induction
- Credibility of assertions
- Identification of assumptions

A reliability analysis was conducted using Cronbach Alpha Coefficient to determine the internal consistency of the subscales and overall scale of the CCTT-X for this sample. The results showed that the overall Cronbach Alpha Coefficients of the test instrument (71 items) are 0.66 and 0.70 for the pretest and posttest respectively. Unlike the posttest reliability estimate, the pretest reliability coefficient is relatively lower than the reliability estimates reported in the test administration manual which

range from 0.67 to 0.90. The low range of the reliability coefficients for the different scales of the test (0.30- 0.67) was expected. According to [19], it is due to “the overlap between sections, the moderate number of items in the parts [especially the last scale, *Assumption Identification*] and the probable heterogeneity of critical thinking”. Due to this, it was therefore concluded that the scales’ reliability estimates for this study were within the acceptable range. However, since the alpha reliability coefficient was rather low for ability to identify assumptions in argument skill (0.30 for both pre-and post-test), the results of analysis for this subscale should be interpreted with caution. Thus, in the following analyses of students’ performance of the CCTT-X subscales, any significant findings were treated with caution.

RESULTS

Results from the Quantitative Data

Research Question One: Are There Any Differences in the Level of Development of Critical Thinking Skills When Peer Evaluation, Self-evaluation and Peer Review Are Adopted?: To find the answer to this question, the ANCOVA test was used to determine the significance of differences between the scores of the CCTT-X. The results of the ANCOVA for post-test with pre-test as covariate showed that the different treatments received by the groups had a marginal effect to their performance in the CCTT-X ($F(3,94)=2.79, p=0.05$). To further investigate how the groups differed, Bonferroni post hoc tests were run. Bonferroni tests adjust the significance level to allow for the multiple testing made when several group comparisons are made using the same data (Hochberg, 1988). The results of the Bonferroni post hoc test based on adjusted post-test means which were calculated at the overall pre-test mean value of 37.91 are as in Table 2.

From Table 2, the largest difference was between the peer review and peer evaluation groups (adjusted mean difference of 3.72) but even this was not significant. This means that, in this study although the students

scored differently in the post-test, the difference in performance level was not statistically significant. Thus, it could not be concluded that one group outperformed others in their general critical thinking ability due to the different treatment received.

A further analysis was then conducted to investigate if there were differences when the four skills in the CCTT-X were analysed individually. For this, the ANCOVA test was used. The results are as displayed in Table 3.

From Table 3, it was revealed that there was a significant effect of treatment for the first two skills which were *inductive reasoning* and *credibility of assertion* ($p=0.01$ and $p=0.02$ respectively) but none for the other two skills namely *deductive reasoning* and *identifying assumptions* ($p=0.24$ and $p=0.70$ respectively). Follow-up tests were then conducted to evaluate pairwise differences among the adjusted post-test means for the *inductive reasoning* and *credibility of assertion* skills scored by each group. The tests analyses for each skill were done by first calculating the adjusted posttest mean scores and then assessing the differences using Bonferroni adjustments. The results yielded from the tests based on the overall pretest mean value of 13.68 and 11.03 for *inductive reasoning* and *credibility of assertions* respectively are simplified in Table 4.

From the results, it was found that the only marked difference in the groups’ performances for the ability to apply *inductive reasoning* was between the peer review and the peer evaluation groups as a significant mean difference was observed between the two groups ($p<0.01$). In particular, the participants in the peer review group significantly outperformed the participants in the peer evaluation group in their ability to apply inductive reasoning with a mean difference of 2.34. No other pairs with significant performance differences were indicated. For the *credibility of assertions*, only the peer review group and self-evaluation group showed a significant mean difference ($p<0.01$). The peer review performed better than self-evaluation group.

Table 2: Post hoc Bonferroni test for pairwise comparisons for CCTT-X

Group vs. Group	Adjusted Group Means	Mean Difference	Sig.
C vs. PR	39.69 vs. 41.04	-1.34	1.00
C vs. SE	39.69 vs. 37.87	1.82	1.00
C vs. PE	39.69 vs. 37.32	2.37	0.54
PR vs. SE	41.04 vs. 37.87	3.17	0.20
PR vs. PE	41.04 vs. 37.32	3.72	0.06
SE vs. PE	37.87 vs. 37.32	0.55	1.00

Table 3: ANCOVA for Treatment and Posttest with Pretest as Covariate (N=99)

Variable	Source of variation	Sum of squares	df	Mean square	F	Sig.
Inductive reasoning	TREATMENT	81.82	3	27.27	4.02	0.01**
	PRETEST	175.39	1	175.39	25.86	0.00
Credibility of assertions	TREATMENT	70.41	3	23.47	3.64	0.02*
	PRETEST	88.17	1	88.17	13.66	0.00
Deductive reasoning	TREATMENT	17.02	3	5.67	1.43	0.24
	PRETEST	215.14	1	215.14	54.12	0.00
Identifying assumptions	TREATMENT	4.78	3	1.59	0.47	0.70
	PRETEST	26.72	1	26.72	7.94	0.01

* $p < 0.05$

** $p < 0.01$

Table 4: Post hoc Bonferroni test for Pairwise Comparisons

Critical Thinking Skill	Group vs. Group	Adjusted Group Means	Mean Difference	Sig.
Inductive Reasoning	PR vs. PA	14.65 vs. 12.31	2.34	0.01**
Credibility of Assertions	PR vs. SA	12.18 vs. 9.79	2.39	0.01**

** $p < 0.01$

Table 5: Pearson's Correlation Coefficients for CCTT Performance and Academic Writing Ability (N=99)

Group	N	Inductive Reasoning		Credibility of Assertions		Deductive Reasoning		Identifying Assumptions		CCTT Overall	
		r	Sig.	r	Sig.	R	Sig.	r	Sig.	r	Sig.
CG	24	0.21	0.34	0.00	1.00	0.10	0.65	0.32	0.13	0.23	0.29
PR	24	0.34	0.11	0.45*	0.03	0.14	0.51	0.16	0.47	0.46*	0.03
SA	24	0.32	0.12	0.35	0.09	0.02	0.93	0.13	0.56	0.38	0.07
PA	27	0.23	0.26	0.35	0.08	0.35	0.08	0.24	0.24	0.42*	0.03

* $p < 0.05$

Research Question Two: What Is the Impact of the Different Treatments on Academic Writing Ability?:

To see the impact of the different treatments on academic writing ability, an investigation of the correlation between the students' critical thinking skills and academic writing ability at the end of the interventions was done. To measure this, the CCTT-X posttest scores were correlated with the students' final term paper scores. Table 5 below displays the results of Pearson's Correlation Coefficients of the two.

Overall, significant correlations between critical thinking skills and academic writing ability were observed on the peer review group and peer evaluation group (both at $p=0.03$). The peer review group even indicated a marked relationship between their academic writing ability and credibility of assertions ($p=0.03$).

All the above results signified that although peer evaluation did foster critical thinking, peer review activity did it better. Thus, further investigation was conducted to unravel what was there in peer review activity which was absent in peer evaluation that led to such results.

Results from the Qualitative Data: Analyses of the oral and the written responses from the interviews and survey questionnaires respectively revealed that all the three experimental groups reported benefitting from the learning experiences. Critical thinking was triggered as reflected in these students' responses.

I have never done this (peer evaluation) before, so I had to think more, "Is it appropriate to use this source?" Then, I looked at the explanation, "Does it match - with the sources she is using?" (PE-21, Interview)

When evaluating, I didn't know whether to put 4, 5. I reread it, "Is she good or not?" I seemed to think. (PE-17, Interview)

You have to explain why this and that (when peer checking). (PR-6, Interview)

When reviewing, we have to review other people's work and we have to think what the best is for that person. We may come with the ideas to help the writer. (PR-23, Survey response)

I think the self evaluation form is really helpful because I can see the differences between my expectations and madam's (the instructor's) expectations. When there are some mismatches, I will improve my term paper. (SE-29, Survey response)

After I have done my term paper, I can see clearly my work and after my instructor gives some corrections and comments, I know where my weaknesses and strengths are. (SE-26, Survey response)

The students benefited from the collaborative learning environment either via peer evaluation or peer review activities as commented by a student. He wrote:

I think that peer advice is also important instead of guidelines by lecturers. At the beginning, I felt that I was weak writing an argumentative essay, but by comparing with others', I know what I also can do. Thus, peer evaluations are important. (PE-13, Survey response)

Besides, the tools used in the activities, that is, the rubric and the checklist provided the students with the criteria guiding their assessment task. The students commented:

If you want to comment just briefly, it's really hard. So, commenting by using those points is much easier because based on the points, I know what to check. (PR-17, Interview)

Listed (in the rubric) were a lot of items - which we needed to have in the writing, didn't we? So, we studied them and we knew that we, "Oh, I don't do well in this,"... like the citation - something was wrong there, so later we corrected it. We tried to get the excellent ranking in the list. (SE-20, Interview)

Although the opportunity to provide critical comments, exchange ideas and defend one's work seems to be more prevalent among the peer review group, it had been hypothesised that the peer evaluation activity would have been able to offer more learning advantages than the peer review activity by including the evaluation feature in order to encourage a more active judgment of the quality of the drafts. Not only did the students question their peers, they actively questioned themselves as they evaluated a peer's work as reflected by student PE-17:

When evaluating, I didn't know whether to put 4, 5. I reread it, "Is she good or not?" I seemed to think. (PE-17, Interview)

As for the self-evaluation group, although the students had the evaluation activity, they did not have the opportunity to discuss and exchange opinions with their peers. This was a contributing factor to the significantly lower performance in the CCTT-X test particularly for *credibility of assertions* as the students lacked the opportunity to practice assertiveness through interactions with others in discussion activities.

One possible reason for the weaker performance of the peer evaluation group despite having the learning activity was the time constraint. It was observed that although the students in the peer evaluation group spent a reasonable time to rate the peers' drafts, they were left with very little time to discuss the comments they had for the peers which they were initially required to do. The time constraint was especially felt when the peer evaluation group had an extra task to do. Not only the students had to read and discuss the drafts, but they also had to assess them. A student from the peer evaluation group pointed this out when asked the difficulty faced in the study. He said:

May be too short (the time given to complete the tasks) - in terms of checking the essay, so we cannot identify more problems." (PE -12, Interview)

The additional time needed by the peer evaluation group to grade their peers' work had reduced the discussion activity which was supposed to be done after each draft was read and rated. The use of the rubric without any discussion activity was insufficient to foster students' critical thinking development. The importance of discussion to foster critical thinking skills development has been reported in the literature [21, 22]. Through discussion, students heighten their thinking as they question, negotiate, defend, clarify and justify ideas.

In addition to constraints of time and discussion opportunity, the multiple tasks of reading, understanding, rating and discussing the drafts could be another explanation for the peer evaluation group lower performance. There was too much to do thus some students failed to (or resolved not to) become actively involved. Some did not have the motivation to do so. The students said:

My friend, he did not really read it. Just roughly – a little bit of assessment – okay and then comment. (PE-13, Interview)

If we only do our job and the other partners don't do their work, it'll be more difficult to share the problem or to share the ideas.” (PE-12, Interview)

DISCUSSION

Students' active engagement in the learning activities is an integral part in developing critical thinking skill. A more thought-provoking activity to engage students in the learning process would be useful for this purpose. Peer evaluation allows formative evaluations for learning in the classroom by giving some authority to peers' role in the process. The exercise allows students “to practice the [metacognitive] skills needed for life-long learning (particularly, evaluation and critical thinking skills) by evaluating others and observing how others evaluate the results of their learning” [23] (p. 2). The peer evaluation interventions in this study provide a hint that the activity can promote critical thinking in academic writing lessons. Among the four critical thinking skills tested in the Cornell Critical Thinking Test [19], the peer evaluation activities helped improve three skills namely credibility of assertions, deductive reasoning and identification of assumptions. Although statistically, peer review activities sparked a higher critical thinking development particularly in the inductive reasoning ability, the finding suggested that the peer evaluation activities, together with sufficient time and discussion opportunities on the writing task among the peers had the potential to make a better impact on students' critical thinking development. A better quality academic writing can then be produced with the help of not only the teacher but also the peers.

CONCLUSION

This study has shown that peer evaluation can foster critical thinking within an academic writing course. Furthermore, it has provided evidence of the importance of collaborative activities within academic writing courses where students are engaged in thinking critically about the quality of their own work and that of their peers. This study suggests that in order to promote critical thinking, teachers should consider using a rubric which will identify the skills being fostered. This requires the students to make judgments on the quality of the work

being evaluated. Teachers should also give adequate time for peer discussion of the rubric ensuring that they and their students understand both the value of critical thinking and how the evaluation and discussion activities foster critical thinking skills.

REFERENCES

1. Vardi, I., 1999. Developing critical writers at the undergraduate level: Some insights from critical thinking pedagogy and linguistics. HERDSA Annual International Conference. Melbourne, Australia.
2. Evans, D., 2008. Reflections on peer evaluation in an English language course. *J Nurs Studies N C N J*, 7(1): 41-49.
3. Ryder, R.J., 1994. Using frames to promote critical writing. *Journal of Reading*, 38(3): 210-218.
4. Wiggins, G.P., 1993. *Assessing Student Performance: Exploring the Purpose and Limits of Testing*. San Francisco: Jossey-Bass.
5. Todd, V. and J.C. Hudson, 2007. Using graded peer evaluation to improve students' writing skills, critical thinking ability and comprehension of material in a Principles of Public Relations course. *Journal of College Teaching and Learning*, 4(10): 39-46.
6. Chu-you, C., 2008. An investigation of peer evaluation in EFL college writing. 25th International Conference of English Teaching and Learning. 2008 International Conference on English Instruction and Evaluation. Retrieved March 17, 2009 from <http://www.ccu.edu.tw/fllcccu/2008EIA/English/C09.pdf>
7. Sadler, D.R., 1989. Formative evaluation and the design of instructional systems, *Instructional Science*, 18: 119-144.
8. Rust, C., M. Price and B. O'Donovan, 2003. Improving students' learning by developing their understanding of evaluation criteria & processes. *Assessment and Evaluation in Higher Education*, 28: 147-164.
9. Cho, K., C.D. Schunn and D. Charney, 2006. Commenting on writing: Typology and perceived helpfulness of comments from novice peer reviews and subject matter experts. *Written Communication*, 23: 260-294.
10. Chaves, J.F., C.M. Baker, J.A. Chaves and M.L. Fisher, 2006. Self, peer and tutor assessments of MSN competencies using the PBL-Evaluator. *Journal of Nursing Education*, 45(1): 25-31.

11. Lindblom-Ylänne, S., H. Pihlajamäki and T. Kotkas, 2006. Self-, peer- and teacher-assessment of student essays. *Active Learning in Higher Education*, 7(1): 51-62.
12. Sahin, S., 2008. An application of peer assessment in higher education. *The Turkish Online Journal of Educational Technology -TOJET*, 7(2): 5-10.
13. Mat Daud, N., N.L. Abu Kassim and N.S. Mat Daud, 2011. Reliability of peer and self-assessment as opposed to teacher assessment. In Nuraihan Mat Daud, Isarji Sarudin, Ainon Jariah Muhamad, Engku Haliza Engku Ibrahim and Khairiah Othman (Eds.), *Language Studies in the Muslim World* (pp: 7-24). Gombak: IIUM Press.
14. Johnson, N. and M. Winterbottom, 2011. Supporting girls' motivation in science: A study of peer- and self-assessment in a girls' only class. *Educational Studies*, 37(4): 391-403.
15. Wenji Qu and Shuyi Yang, 2010. A peer- and self-assessment project implemented in practical group work. *Journal of Language Teaching and Research*, 1(6): 776-781.
16. Willey, K. and A. Gardner, 2010. Investigating the capacity of self and peer assessment activities to engage students and promote learning. *European Journal of Engineering Education*, 35(4): 429-443.
17. Langan, A.M., D.M. Shuker, W.R. Cullen, D. Penney, R.F. Preziosi and C.P. Wheater, 2008. Relationships between student characteristics and self-, peer and tutor evaluation of oral presentations. *Assessment and Evaluation in Higher Education*, 33(2): 179-190.
18. Rollins, T.J., 1990. Levels of critical thinking of secondary agriculture students, *Journal of Agricultural Education*, Fall, pp: 47-53.
19. Ennis, R.H., J. Millman and T.N. Tomko, 2005. *Cornell critical thinking tests: Administration manual*. Seaside, CA: The Critical Thinking Co.
20. Hochberg, Y., 1988. A sharper Bonferonni procedure for multiple tests of significance. *Biometrika*, 75(4): 800-802.
21. Totten, S., T. Sills, A. Digby and P. Russ, 1991. *Cooperative learning: A guide to research*. New York: Garland.
22. Tsui, C., 2003. Fostering critical thinking through effective pedagogy. *The Journal of Higher Education*, 73(6): 740-763.
23. Omelicheva, M., 2004. Self and peer evaluation in undergraduate education: Are promises worth risking the perils? Paper presented at the annual meeting of the APSA Teaching and Learning Conference. NA, Washington, DC.