

# ASSESSING CRITICAL THINKING PERFORMANCE OF POSTGRADUATE STUDENTS IN THREADED DISCUSSIONS

Cheng Lee Tan and Lee Luan Ng  
*Faculty of Languages and Linguistics, University Malaya*

## ABSTRACT

Critical thinking has increasingly been seen as one of the important attributes where human capital is concerned and in line with this recognition, the tertiary educational institutions worldwide are putting more effort into designing courses that produce university leavers who are critical thinkers. This study aims to investigate the critical thinking ability of the postgraduate students in threaded discussion. Participants of the threaded discussion were the postgraduate students of a faculty of a public university located in Klang Valley. Data was collected from the postgraduate courses' threaded discussion assignment mediated via a learning management. Overall, four threaded discussion transcripts were collected and used as the source of data of this study. The researcher assessed the four threaded discussion transcripts based Newman et.al (1995) content analysis framework. Results showed that the participants' postings mainly reflected their critical thinking ability in terms of being able to include relevant (R+ positive critical thinking indicator), clear (AC+ positive critical thinking indicator), novel (N+ positive critical thinking indicator) and justified (JS+ positive critical thinking indicator) input into the threaded discussion. It was found that in the postings, the participants integrated a lot of their personal experience into the discussion (O+ positive critical thinking indicator) and were able to link their ideas coherently (L+ positive critical thinking indicator). However, it was found that the participants generally lacked the ability to critically evaluate their peers' or their own postings.

## KEYWORDS

Critical thinking, threaded discussion, computer mediated higher education instruction.

## 1. INTRODUCTION

Critical thinking has increasingly been seen as an important attribute in empowering human capital. For instance, as stated in 10<sup>th</sup> Malaysian Plan, "the success of the innovation agenda hinges on a Malaysian citizenry that values openness, embraces **critical thinking** and encourages risk taking and experimentation. This will require an education system that nurtures creative and analytical human capital" (Tenth Malaysia Plan, 2011-2015). Besides that, as quoted by Koo, Wong, Kemboja, Chang and Mohd Subakir (2011) in their study, Malaysia Ministry of Higher Education has established among others the National Higher Educational Plan which "aims squarely on holistic human capital development, to produce Malaysians who are intellectually active, creative, innovative, adapted and capable of **critical thinking**" in order to address the unemployment situation among public universities graduates. Thus, it is evident that higher educational institutions in Malaysia are encouraged to produce university leavers who possess the capability to think critically in order to help them to secure employment. This perception also prompts the tertiary education institutions worldwide to utilize means such as Computer Supported Collaborative Learning (CSCL) platform when designing teaching and learning activities that encourage critical thinking among the university students.

Since past studies had proven that CSCL can be useful in terms of nurturing critical thinking virtual communities (Li, 2010; Lim, Cheung, & Hew, 2011), threaded discussion which is known as one of the Computer Mediated Communication (CMC) tools have been employed by the instructors in carrying out CSCL activities with their students. To date, a number of studies have investigated the capabilities of asynchronous threaded discussion in cultivating critical thinking among learners (Leston-Bandeira, 2009; Irfan & Hazita, 2010).

In addition, there are numerous studies which investigated the linguistic features found in computer mediated communication transcripts. Despite that, there has been no direct study which investigates how the use of certain linguistic features reflects the critical thinking. In view of what was stated by Leston-Bandeira (2009), which indicated that there is potential in developing critical thinking skills among learners via threaded discussions, so it would be logical to assume the trend of participants in using certain cohesive devices in influencing the quality of postings, and in enabling researchers to gain some perspectives on the participants' critical thinking performance. Thus, this research aims to fill this gap by investigating how the use of cohesive devices such as conjunction could reflect the presence of positive and negative critical thinking indicators in threaded discussions. The two specific objectives of this study are firstly, to determine the students' critical thinking abilities by assessing the threaded discussion transcripts based on content analysis scheme established by Newman, Webb and Cochrane (1995) and secondly, to delve into how the use of cohesive devices such as conjunction in reflecting the critical thinking performance of participants in threaded discussions. In relation to the stated objectives, the following are the research questions.

- 1) What is the frequency and percentage of each of the positive and negative critical thinking indicators attained by each threaded discussion?

## **2. METHOD**

### **2.1 Participants and Setting**

Via the use of a mixed-method approach, the threaded discussion data was collected from one of the public universities located in Klang Valley, Malaysia. Participants of the asynchronous online discussion were the postgraduate students of the Faculty of Languages and Linguistics of that public university. The participants consist of two groups of postgraduate students who were enrolled in two postgraduate courses. One group of the participants were the students who enrolled in the Research Methodology (RM) course while the other group consisted of participants who enrolled in the Second Language Acquisition (SLA) course. The threaded discussion activities form part of the course continuous assessment activities.

### **2.2. Data Collection and Data Analysis Procedures**

Overall, a total of four topics of discussion which are relevant to the course content were assigned during the semester. The topics of discussion were known as RM 1, RM 2, SLA 1 and SLA 2. All these four topics were formulated in line with the goal to fulfill the learning outcomes of the courses. Two topics of discussion were designed for each course. The topics of discussion were posted on the class learning management system. Participants were required to post their personal comment at least once for each question and respond to their coursemates' responses at least twice. However, no limitation was set on the maximum number of responses one could post. The duration for each question to be discussed in the threaded discussion was about two weeks.

After all the online discussion sessions had ended, the online asynchronous transcripts were downloaded and imported into Nvivo 9, a computer data analysis software. In order to answer research question 1, the data were analysed based on Newman et. al (1995) content analysis framework which consists of positive and negative critical thinking indicators. One of the benefits of employing Newman et.al. (1995) content analysis scheme in assessing critical thinking found in threaded discussion transcripts is, it allows critical thinking to be quantified. Newman et.al (1995) content analysis scheme was chosen to analyse the data because it provides an explicit list of positive and negative critical thinking. Being able to quantify critical thinking aids educators or researchers who wish to access their students' critical thinking performance. This is because numerical values obtained could be used as learning evidence that can inform both the educators, researchers and students regarding critical thinking performance.

In order to obtain interrater reliability reading, both the researcher and a second coder coded about 20% of the overall data, and the interrater reliability values for both Kappa Coefficient and percentage of agreement were calculated using the coding compound query feature of Nvivo 9. It was found that the Kappa Coefficient value was 0.72 while the percentage of agreement value was 94.66%.

In terms of unit of analysis, the researcher of this study selected sentence as the unit of analysis. The reason being the use of sentence as unit of analysis was claimed to help in attaining higher interrater reliability Gorsky et.al.(2012). In the next section, the results and discussion will be presented.

### 3. RESULT

Table 1 illustrates the frequency and percentage of each of the positive critical thinking indicator sub categories and Table 2 illustrates the frequency and percentage of each of the negative critical thinking indicator sub categories which belong to the four threaded discussions.

Table 1. The frequency and percentage of positive critical thinking indicator sub categories of the rm 1, rm 2, sla 1, sla 2 threaded discussions

Threaded Discussion Transcripts	Positive Critical Thinking Indicators	A+	C+	I+	JS+	L+	N+	O+	P+	R+	W+	Total
a (i) RM 1	Frequency	213	58	112	72	67	114	46	26	213	60	981
a (ii) RM 1	Percentage %	18.65	5.08	9.81	6.30	5.87	9.98	4.03	2.28	18.65	5.25	85.90
a (iii) RM 2	Frequency	153	29	85	72	124	115	29	48	153	28	836
a (iv) RM 2	Percentage %	17.31	3.28	9.62	8.15	14.03	13.01	3.28	5.43	17.31	3.17	94.59
a (v) SLA 1	Frequency	287	81	40	140	134	247	133	16	287	120	1485
a (vi) SLA 1	Percentage %	18.84	5.32	2.63	9.19	8.80	16.21	8.73	1.05	18.84	7.88	97.49
a (vii) SLA 2	Frequency	130	20	29	45	55	91	45	37	130	54	636
a (viii) SLA 2	Percentage %	19.58	3.01	4.37	6.78	8.28	13.71	6.78	5.57	19.38	8.13	95.79

Table 2. The frequency and percentage of negative critical thinking indicator sub categories of the rm 1, rm 2, sla 1 and sla 2 threaded discussions

Threaded Discussion Transcripts	Negative Critical Thinking Indicators	A-	C-	I-	JS-	L-	N-	O-	P-	R-	W-	Total
b(i) RM 1	Frequency	37	25	15	11	7	19	0	7	31	9	161
b(ii) RM 1	Percentage %	3.24	2.19	1.31	0.96	0.62	1.66	0	0.62	2.72	0.79	14.11
b(iii) RM 2	Frequency	4	6	6	3	10	11	0	0	5	3	48
b(iv) RM 2	Percentage %	0.45	0.68	0.68	0.34	1.13	1.24	0	0	0.57	0.34	5.43
b(v) SLA1	Frequency	5	3	5	6	5	5	0	0	4	6	38
b(vi) SLA 1	Percentage %	0.33	0.20	0.33	0.39	0.33	0.33	0	0	0.26	0.39	2.56
b(vii) SLA 2	Frequency	4	2	1	3	4	2	0	5	4	3	28
b(viii) SLA 2	Percentage %	0.60	0.30	0.15	0.45	0.60	0.30	0	0.75	0.60	0.45	4.20

With reference to a (ii) RM 1 and b (ii) RM 1 presented in Table 1 and Table 2, it is obvious that the overall percentage of positive criticalness (85.9%) is higher than the percentage of negative criticalness (14.11%) and based on the postings found in (i) RM 1 and b (i) RM1, there are 981 occurrences of positive criticalness and 161 occurrences of negative criticalness identified from the transcript.

a (ii) RM 1 shows that R+ (Relevant statement)(18.65%), A+ (Clear, unambiguous statement)(18.65%), N+ (Novelty) (9.98%) and I+ (Important statement) (9.81%) are the most frequently detected indicators and the least detected indicators are P+ (Discussing the practicality of new ideas and suggesting solutions) (2.28%), followed by O+(Referring to outside knowledge/experience)(4.03%) and C+(Critical assessment of others' or own contribution)(5.08%). Where negative criticalness is concerned, b (ii) RM 1 showed that A- (Confuse statements) (3.24), R- (Irrelevant statement) (2.72) and C- (Uncritical acceptance or unreasoned rejection) (2.19) and N- (Repeating what has been said, or false or trivial leads, or accepting first offered solution) (1.66) are the most occurring uncritical thinking traits.

With reference to a (iv) RM 2 and b (iv) RM 2 shown in Table 1 and Table 2, it is found that the positive critical thinking indicators contribute 94.59% while the negative critical thinking indicators contribute 5.43 % to the overall scoring. Based on a (iii) RM 2 and b (iii) RM 3, there are 836 occurrences of positive criticalness and 48 occurrences of negative criticalness.

Referring to a (iv) RM 2 above, A+ (17.31%), R+ (17.31%), L+ (Linking of ideas and generating new data from information collected) (14.03%) and N+ (13.01%) are the most frequently detected positive critical thinking indicators in participants' postings. According to b(iv) RM 2, the most frequently detected negative critical thinking indicators are N- (1.24 %), L- (Repeating information without making inferences or offering an interpretation, or stating that one shares the ideas or opinions stated without taking these further or adding any personal comments ) (1.13 %), C- (0.68 %).

Compared with the findings a (ii) RM 1 and b (iv) RM 2, the overall percentage of negative critical thinking instances in RM2 is much lower. This could be due to the fact that they had received feedback from the instructor on their RM 1 threaded discussion session, causing the participants to know how to engage in the threaded discussion and thus perform better. Apart from the instructor's feedback, topic familiarity was likely to cause the participants to perform better in RM 2 threaded discussion since Chou and Chen (2010) discovered that topic familiarity could affect their subjects' willingness to write postings. Thus, it seems imperative for the participants to have sufficient background knowledge towards a particular topic of discussion in order to ensure they are familiar with the topics of discussion because it would encourage them to write and post postings.

As presented in a (vi) SLA 1 and b (vi) SLA 1 respectively, it is obvious that, the overall percentage of positive criticalness (97.49%) is higher than the percentage of negative criticalness (2.56 %). According to a (v) SLA 1 and b(v) SLA 1 shown in Table 1 and Table 2, there were 1485 occurrences of positive criticalness and 38 occurrences of negative criticalness detected from the transcript.

Where positive criticalness of the SLA 1 threaded discussion is concerned, R+ and A+ which shared the same percentage value that is 18.84%, N+ (16.21%) and JS+ (Justification) (9.19%) are apparently the four most frequently detected indicators. These findings indicate that the participants were able to produce relevant and clear statements, bring in novel ideas and rationalize their arguments. However, their lack of competence in incorporating important input and critiquing their peers' contributions was detected if one refers to the percentage of both I+ and C+. The P+ indicator's percentage value is the lowest. This was probably because the topic of discussion was not designed for problem solving and it did not require participants to suggest solutions and discuss the practicality of each solution.

Referring to b (vi) SLA 1 shown in Table 2, it seems that W- (Narrowing the discussion) and JS- (Unjustified statement) are the two most frequently found negative critical thinking indicators which have the same percentage value that is 0.39%. They were followed by A- , L-, I- and N- indicators which also share the same value that is 0.33%.

Based on a (viii) SLA 2 and b (viii) SLA 2, it is noticeable that the percentage of positive criticalness (95.79%) is higher than the percentage of negative criticalness (4.2 %). In addition, a (vii) SLA 2 and b(vii) SLA 2 showed that there were 636 occurrences of positive criticalness and 28 occurrences of negative criticalness spotted in the transcript.

Where the positive criticalness of SLA 2 threaded discussion is concerned (See a (viii) SLA 2), both R+ and A+ indicators show the same percentage value that is 19.58%, followed by N+ indicator at 13.71% and L+ indicator at 8.28%, rendering the four of them to be the four most scored indicators while the least scored indicators are C+ at 3.01%, followed by I+ at 4.37% and P+ at 5.57%. These findings indicate that the

participants were able to produce relevant and clear statements, bring in novel ideas and link ideas and create new interpretation. However, similar with the findings in a (v) SLA 1, their lack of competence in incorporating important input, suggesting and providing solutions and critiquing their peers' contributions contributed to the low frequency counts of P+, I+ and C+ indicators. This is probably because the topic of discussion was not designed for problem solving and it did not require participants to suggest solutions and discuss the practicality of each solution. In addition, the low percentage value of C+ also suggested that the participants did not engage much in reflecting and commenting on their peers' comments critically. They simply agreed to what others said without further probing or providing reasons for their stance of agreeing to their peers' postings.

Referring to b (viii) SLA 2, it is noted that where the negative criticalness of the SLA 2 threaded discussion is concerned, the percentage values of the four most frequently detected negative critical thinking indicators found in SLA 2 threaded discussion transcript the were P- at 0.75%, L- at 0.60%, A- at 0.60%, and R- at 0.60%. As compared with the results tabulated in b (vi) SLA 1, generally, the similarity found between the coding for the SLA 1 and SLA 2 is there is a zero percentage of O-. The overall percentage of negative critical thinking indicators is also found to be higher in SLA 2 than in SLA 1. The lack of interest, followed by the lack of motivation to participate in threaded discussion, may be also the reasons why the participants overall critical thinking performance dropped in the SLA 2 threaded discussion as compared to the SLA 1 threaded discussion.

#### 4. DISCUSSION

After inspecting , it is found that all the participants seemed to have less problems in contributing inputs that were assigned codes R+, A+, N+, JS+, L+ (Linking of ideas and generating new data from information collected) and O+. Interestingly, it is also noted that none of O- (Squashing attempts to bring in outside knowledge or sticking to prejudice or assumptions) input is detected in the four threaded discussion transcripts. Both the O+ percentage (Referring to outside knowledge/experience) for the SLA 1 and SLA 2 threaded discussions are also higher than that of the RM 1 and RM 2 threaded discussions. This may be due to the nature of topic.

The characteristics of both the SLA 1 and SLA 2 topics of threaded discussion seem to encourage the participants to bring in more of their own previous experience and background knowledge and also draw in more related outside materials to substantiate their arguments. The participants also proved that they were able to integrate their previous experience and background knowledge into the threaded discussion. This is probably because most of the participants were exposed to the second language acquisition theories during their undergraduate study. Therefore, it was easier for them to rely on their background knowledge and experience and include other relevant materials they gained from books, articles and internet into both the SLA 1 and SLA 2 threaded discussions.

On the other hand, in the case of the participants who participated in the threaded discussion sessions of the course Research Methodology, these students seemed to exhibit behaviour that showed that they did not have enough experience pertaining to relevant aspects of research which they were required to discuss. These students also did not show much abilities in locating and inserting external relevant information into the content of the discussion. This may be due to their lack of previous exposure to aspects of research, or their failure to relate to the existing knowledge. Another probable explanation may be linked to the fact that the research methodology course is an introductory course which is designed to expose the postgraduate students who are relatively new to the important concepts of academic research.

The importance of being able to locate and insert relevant information, be it taken from the outside materials such as books, personal experience or previous knowledge into the threaded discussion was highlighted by Woo and Wang (2009) study. Woo and Wang (2009) conducted a study to find out whether web blogging was effective in encouraging critical thinking. They were also interested in investigating the influence of different kind of blogging topic has on the overall results of the frequency of each critical thinking indicator. Their participants were secondary school students. The three topics of web blogging were designed based on secondary school History subject syllabus. The scheme they used to code their web blogging transcripts was similar to the one employed by the current researcher which is the Newman et.al (1995) content analysis scheme. They designed three topics for students to discuss in total. The similarity

found among the three threaded discussions was that R+, O+ and JS+ indicators were the three most frequent detected critical thinking indicators in all the three web blogging transcripts. They also claimed that because the topic 1 and topic 3, which allowed the students to use the information from the textbooks to substantiate their arguments, the C- percentage reported was insignificant as it was too small a value. On the other hand, topic 2 which was designed in such a way that the participants could not rely solely on the textbook information to support their arguments, the overall negative criticalness of topic 2 increased and C- indicator percentage became significant. The same phenomenon is observed in the findings of this present study. In the data of this study, it is noted that while the percentage of O+ (Bringing in outside knowledge/experience to bear on the problem) indicator in both the SLA 1 and SLA 2 threaded discussions are higher than that of the RM 1 and RM 2 threaded discussions, the percentage of C- (See Table 2) indicator for both the SLA 1 and SLA 2 threaded discussions are also lower than that of the RM 1 and RM 2 threaded discussions. These findings again show that being able to include relevant outside materials into the threaded discussion could be a crucial key to improve the participants' overall critical thinking performance.

For this study, another similarity identified throughout the four topics of threaded discussion transcripts was that the C+ indicator, which was recognized as one of the least detected indicators in all the four threaded discussion transcripts. This may suggest that in general participants did not engage much in evaluating their peers' postings critically. The same phenomenon was also observed and reported by Irfan and Noor (2010) who examined the trainee teachers' online discussion forum transcripts with the aim of finding out the positive critical thinking and negative critical thinking indicators that were exhibited. As mentioned by Irfan and Noor (2010), the subjects of their study were the trainee teachers who were all novice teachers. In other words, they did not have any teaching experience prior to their teaching practice. The online discussion forum was launched during their teaching practice in order for them to interact with their peers and lecturer by sharing and discussing the problems they encountered during teaching practice. One of their findings was that C+ indicator was one of the six least found positive critical thinking indicators, rendering them to assert that their subjects of study might be weak in their ability to evaluate their peers' and their own postings critically and the researchers posited that this was probably caused by their lack of real life teaching experience.

## 5. CONCLUSION

Critical thinking can be cultivated through engaging students in computer supported collaborative learning activities such as threaded discussion. The postings from threaded discussions can be assessed in terms of critical thinking ability. In this study, based on Newman et. al (1995) content analysis framework. it seems that the R+ (Relevant statements), A+ (Clear and unambiguous statements), N+ (Novelty), JS+ (Justified statements), L+ (Linking of ideas and generating new data from information collected) and O+ (Referring to outside knowledge/experience) were the most frequently detected indicators for all the four transcripts of threaded discussions. On the other hand, C+ (Critical assessment of others' or own contribution) was one of the hardly detected indicators found in all the four transcripts of threaded discussions. In addition, where the percentage of O+ (Bringing in outside knowledge/experience to bear on the problem) is higher, the percentage of C- (Uncritical acceptance or unreasoned rejection) seems to be lower. The results also indicate that participants who had sufficient amount of knowledge and personal experience pertaining to the matters under discussions, would most likely rationalize their responses in postings, and not exhibit behaviours that reject or accept others' viewpoints uncritically. Therefore, this may suggest that the amount of the personal experience and knowledge the participants have towards a particular topic of threaded discussion could affect their critical thinking performance. In short, participating in a well-designed threaded discussion task can inculcate critical thinking as it prompts the participants to share and reflect on the issues pertaining to course content.

In the future, the instructors may want to consider assessing the content of students' assignments in terms of critical thinking performance by employing rubrics that are designed specifically for assessing critical thinking. One of the means that may assist the instructors to detect the critical thinking aspects of students' work can be the use of cohesive devices such as the use of conjunctions and collocations. The instructors perhaps can draw the attention of students to the use of linguistic elements, particularly the use of cohesive devices in facilitating the flow of ideas. For instance, the instructors can demonstrate how the use of

conjunction can impact the writing quality by comparing writing samples. The instructor can also show the students the various manners they can employ the first personal pronoun 'I' in conveying personal but substantiated comments and stance. The instructors can also do the same thing when it comes to the use of substitution and ellipsis. Helping the students to acquire the skills of using linguistic elements such as conjunction, pronoun 'I', substitution and ellipsis effectively in the threaded discussion context is crucial as these skills help them to argue logically and improve their critical thinking ability. This can then act as a complement to the conventional rubrics which are used to measure the writing quality of postgraduate students' assignments.

## REFERENCES

- Chou, P. N., & Chen, W. F. (2010). Chinese students' perceptions of online learning in Western discussion boards: A cultural perspective. *International Journal of Instructional Technology & Distance Learning*, 7(2), 35-48.
- Halliday, M. A. K., & Hasan, R. (1976). *Cohesion in English*. London: Longman
- Gorsky, P., Caspi, A., Blau, I., Vine, Y., & Billet, A. (2012). The impact of unit (sentence vs message) on the results of quantitative content analysis. *The International Review of Research in Open and Distance Learning*, 13(1).
- Irfan, N. U., & Noor Hazita, A. (2010). Trainee teachers' critical thinking in an online discussion forum. *MJLI*, 7, 75-91.
- Koo, Y. L., Wong, F. F., Kemboja, I., Chang, P. K., & Mohd Subakir, M. Y. (2011). Literasi kritikal dalam konteks pendidikan tinggi: Suara dari sebuah bilik darjah di Malaysia. *GEMA Online™ Journal of Language Studies*, 11(2), 21.
- Leston-Bandeira, C. (2009). Using e-learning to promote critical thinking in politics. *Enhancing Learning in the Social Sciences*, 1(3).
- Li, K. M. (2010). Integrating weblogs in a pedagogy model for enhancing students' critical thinking skills. *Research and Practice in Technology Enhanced Learning*, 5(1), 35-49.
- Lim, S. C., Cheung, R. W. S., & Hew, K. F. (2011). Critical thinking in asynchronous online discussion: An investigation of student facilitation techniques. *New Horizons in Education*, 59(1).
- Newman, D. R., Webb, B., & Cochrane, C. (1995). A content analysis method to measure critical thinking in face-to-face and computer supported group learning. *Interpersonal Computing and Technology*, 3(2), 56-77
- Tenth Malaysia Plan 2011-2015. Retrieved from [https://www.pmo.gov.my/dokumenattached/RMK/RMK10\\_E.pdf](https://www.pmo.gov.my/dokumenattached/RMK/RMK10_E.pdf)
- Wang, Q. Y., & Woo, H. L. (2009). Using weblog to promote critical thinking-An exploratory study. *International Journal of Human and Social Sciences*, 872-880. doi: 10.1007/s12564-010-9101-5