

# *Critical Thinking & Reading Skills: A Comparative Study of the Reader Response & the Philosophy for Children Approaches*

MOOMALA OTHMAN & ROSNANI HASHIM

## Introduction

Teaching critical thinking in schools has become a necessity, not only in Malaysia, where the research reported here took place, but elsewhere in the world. Teaching students to think critically has a very high premium in education because it is considered to be directly related to cognitive development. It is understood to enhance ability in reasoning skills, logical skills and also in reading and mathematics (Lipman, 1980; Reed, 1992). It has been universally argued that critical thinking ensures one's success academically and professionally (Abbott, 1997; Gelberg, 1993; Resnick, 1989).

Since endorsing this view of the importance of critical thinking for students' overall development, Malaysia has incorporated, in stages, critical thinking skills into schools since 1989 (Suhailah, 2001). This is in line with the Malaysian educational policy termed the National Education of Philosophy (NEP). The NEP emphasizes holistic education designed to produce an intellectually, spiritually, physically and emotionally balanced character. As such, philosophy acts as a framework for curriculum design and its practices. To promote intellectual development, the Ministry of Education has emphasized student-centered teaching methodologies such as the inquiry and discovery method, the Socratic method, discussion in general, and project and group work. To promote critical thinking skills, the strategy adopted is an infusion approach, whereby the teaching of thinking skills

is incorporated in all lessons or subjects taught, including the subject of English, which is, of course, in Malaysia a second—if widely used—language.

However, a review of the literature indicates that the infusion approach is fraught with uncertainties. Its success is dependent on many factors, including the attitude and behavior of teachers, teacher preparation, and constraint factors such as crowded syllabuses and shortage of time. Thus, Suhailah (2001) has suggested implementing other programs that have been successful in teaching thinking skills to students. This study aims to explore two of those—Philosophy for Children and the Reader Response program—and to assess their relative success in this regard.

The Reader Response approach is based on a literary criticism model, and is widely used in language and literature programs in the schools. In this approach, the reader plays a prominent role in interpreting a text: she is seen as an active participant in the creation of the meaning of the text read. This departs from a traditionalist view, which asserts that the interpretation of a piece of literature lies

---

*Moomala Othman (rozmalasan@yahoo.com) is Senior English Language Teacher, Gombak Integrated Residential School at Sg Pusu, and is affiliated with the Centre for Education and Human Development at International Islamic University Malaysia. Rosnani Hashim (rosnani@iiu.edu.my) is Professor of Education at the Institute of Education, International Islamic University, Malaysia (IIUM), and author of Educational Dualism in Malaysia (Oxford University Press, 1996) and The Teaching of Thinking in Malaysia (IIUM, 2003).*



within the text. On the contrary, the Reader Response approach proposes that the interpretation of a text is dependent on the interaction of the reader's background knowledge and the text (the author's interpretation). The resulting "response" is a dialogic product of the text and the reader. The process of reading leads to discussion and further reasoning, resulting in the creation of different versions of meaning. Thus the program presumes to teach critical thinking through the reasoning process and discussion which accompany learning how to critique a piece of literature.

The Philosophy for Children program (P4C) is a thinking program which also uses stories—stories specifically written to raise important philosophical issues in the readers' minds, which then form the topics of subsequent discussion. The P4C novels are read in a group, usually aloud, and the Socratic conversation which follows explores various interpretations of the meaning of the text as well as the philosophical issues which the text has awakened for the readers.

It is interesting to note that the methodology adopted in the Reader Response approach as well as the Philosophy for Children (P4C) approach are those recommended by critical thinking proponents. Some of the strategies recommended are: discussions of controversial issues; collaborative learning; metacognition; questioning strategies (Socratic questioning, inductive questioning or deductive questioning); the use of content-based language termed immersion; relating or finding relevance in terms of what students learn in the outside world so that they find personal meaning; and lastly, learning strategies. In the Reader Response approach, the methodology adopted includes questioning (Page, 2001), brainstorming, journal writing, the use of literature logs, group discussion and/or responding to their peers' opinions or responses, role play and displaying students' writing or oral response (Miller, 2002). P4C, on the other hand, stresses discussion, dialogue, Socratic questioning, responding to peers' opinions, collaborative learning, reasoning, and debating (Lipman, 1993).

Both approaches use reading materials whereby students give various interpretation of the meaning of the text read. This calls for active participation in the process of learning or giving opinions or interpretations or meaning of what is read. Since both approaches involve reading, the study also looked at the relationship between critical thinking and reading skills. At the same time it aimed to find which approach would improve students' reading skills, as well as accommodate a reading-for-meaning model.

Active involvement is essential to both approaches. Via this active involvement, the thinking that takes place is made explicit and critical thinking is promoted and maximized. This agrees with the notion that critical thinking involves participants daring to take risks in voicing their opinions or interpretations on important issues in ongoing conversations (Nussbaum, 2002)—for, as Langer (1997) argues, the methods of instruction in the classroom

have a direct effect on the process of learning and thinking that takes place among learners.

### Objectives of the Research

The general objective of this study was to explore the strengths and limitations of two different teaching approaches in enhancing critical thinking in the English classroom in Malaysia. It intended to determine if the two approaches improve students' critical thinking, and to assess which of the two approaches is more effective in this regard. Since reading is involved in both approaches, this study also looked at the effect of each approach on reading skills, and, subsequently, the relationship between critical thinking and reading skills. Thus the study aimed to answer the following questions:

1. Is there a statistically significant difference in means of the pre and post tests for critical thinking skills using the New Jersey Test of Reasoning Skills (NJTRS) for each treatment group?
2. Is there a statistically significant difference between the Reader Response (RR) and the Philosophy for Children Program (P4C) groups for critical thinking?
3. Is there a statistically significant difference in means of the pre and post test for reading skills using TOEFL for each treatment group?
4. Is there a statistically significant difference in means between the Reader Response (RR) and the Philosophy for Children Program (P4C) for reading skills?
5. Is there a correlation between the mean scores of the NJTRS and the mean scores of the reading skill across the two groups?

### Methodology

The school selected for this study was a fully residential science school in a rural setting in Malaysia. The population was homogenous, comprising Malay students who had achieved good grades in the standard government examinations, for instance, the PMR Examination. Most of the students were from the state of Selangor or Wilayah Persekutuan and were from middle to upper income family groups. The research design adopted was an experimental one, which aimed to find out which of these two approaches or treatments was better at fostering critical thinking and reading skills.

The population of the form fours (ages 12-13) in the school was about 125 students and they were assigned in five classes through convenience sampling. Subsequently, to carry out this experimental design, two classes out of the five were randomly assigned as the treatment group one and two respectively. Treatment group one was the P4C group, and treatment group two was the RR group. Each group had 24 students. Pre and post tests were given to both groups, resulting in quantitative data. Qualitative data were also gathered in the form of students' journal entries, observation during the lessons, and teachers' journal entries. Before



collecting the data, a pilot test was carried out to determine the reliability of the instruments used. The experimental study involved three stages. Stage one was the pre-test stage, where the critical thinking and reading test was conducted. The second stage involved giving the respective treatment, which was carried out for 16 weeks. The third stage was the post-test stage, where the two tests were carried out again.

From the tests scores, the *t* test was run to determine if there was a statistically significance mean difference between pre and post tests for each treatment and between the two treatments. For the *t* test, the significance level was set at  $p < 0.05$ . In analyzing the data, a triangulation strategy employing qualitative instruments was employed. Triangulation was also used to sustain findings or conclusion drawn from the quantitative analysis.

### Instrumentation

The P4C reading text chosen for use was *Lisa* (Lipman, 1983), used for grades 7 – 9 in the United States This novel focuses upon ethical and social issues such as fairness, lying and truth telling. Other issues explored include the rights of children, job and sex discrimination, and animals' rights. The book comes together with a manual (Lipman, 1983) in which are compiled the activities that could be carried out or issues that could be discussed. The Reader Response (RR) group used the Literature text recommended by the Malaysian Ministry of Education (2000): *Selected Poems and Short Stories for Form 4 Literature in English for Upper Secondary Schools*.

The instrument used to measure reasoning skills was the New Jersey Test of Reasoning Skills (NJTRS). This test was developed by Dr. Virginia Shipman, then a Senior Research Psychologist at the Educational Testing Service in Princeton, New Jersey (Shipman, 1983). The test had 50 multiple-choice items, representing 22 reasoning skills areas. It is a test of ability to reason, rather than a test on scientific inquiry or on judgment. It is clearly and simply written (its Flesch reading level is 4.5) and its reliability (ranging from 0.84 to 0.91) compares favorably with other thinking tests such as the Cornell Critical Thinking and the Whimbey Analytical Skills program. Since the subjects involved were upper secondary forms, an adopted version of TOEFL was used to measure reading skill.

### Results and Discussion

Inferential statistics were employed to investigate if the two groups were comparable in critical thinking and reading skills. Table 1 shows the results of the independent *t* test for means of the two groups on the pre test scores of the critical thinking ability (NJTRS scores). There was no statistical significant difference (Table 1) in the pre test scores of both groups. This showed that the two groups were comparable in their critical thinking skills at the beginning of the experiment, although RR has a higher mean than the P4C group.

**Table 1:** Two-tailed *t* test between the treatment groups for the pre-test NJTRS scores.

Treatment Groups	N	Mean	M.dif f	S.d.	t	df	P
P4C	22	33.68	-1.32	5.37	.917	41	.365
RR	22	35.00		3.91			

Results of the independent *t* test on the reading skill ability also revealed that the two groups were comparable in their reading skills at the beginning of the experiment. As shown in the following table (Table 2) there was no significant statistical difference.

**Table 2:** Two-tailed *t* test between the treatment groups for the pre-test TOEFL scores.

Treatment Groups	N	Mean	M.dif f	S.d.	t	df	P
P4C	20	35.22	0.76	5.27	4.28	43	.671
RR	20	35.45		6.6			

To investigate the research questions involved in this study, inferential statistics were also employed.

The result of a paired difference *t* test for the pre and post tests (see Table 3) showed that there was a statistically significant mean difference for the reasoning skills in the P4C group, ( $t = .993$ ,  $df = 21$ ,  $p < 0.025$ ). Therefore, there was adequate evidence to reject the null hypothesis ( $H_0$ ).

**Table 3:** The paired difference *t* test for the critical thinking skills of the P4C group.

	Mean	M.dif f	N	S.d.	t	df	P
Pre NJTRS	33.68	2.27	22	5.37	-2.993	21	.007*
Post NJTRS	35.95		22	5.21			

\* Significant  $\alpha = 0.025$

On the other hand, Table 4 indicates that, there was adequate evidence not to reject the null hypothesis ( $H_0$ ) for the Reader Response group, ( $t = -.929$ ,  $df = 19$ ,  $p > 0.025$ ).



**Table 4:** The paired difference *t* test for critical thinking skills of the RR group

	Mean	M.dif f	N	S.d.	t	df	P
Pre NJTRS	34.90	0.5	20	3.99	-.428	19	.638
Post NJTRS	35.40		20	4.82			

Thus, it can be concluded that there was no statistically significant mean difference in the pre and post tests in the critical thinking skills in the RR group. Only the P4C group showed a significant difference in the improvement of critical thinking skills. The RR group also showed enhanced critical thinking skills (mean diff 0.5), but the improvement was not statistically significant.

An independent two-tailed *t* test was employed to compare the two groups of the P4C and the RR for critical thinking. It was also to determine if there were significant differences in the critical thinking scores after undergoing the treatment of the two approaches.

**Table 5:** Two-tailed *t* test between the treatment groups for critical thinking.

Treat- ment Groups	N	Mean	M.dif f	S.d.	t	df	P
P4C	22	35.954	0.5	5.214	.336	42	.738
RR	22	35.454		4.626			

The results of the *t* test (see Table 5) reveal that the mean score in critical thinking was higher for the P4C group than for the RR group. However, the null hypothesis was not rejected,  $t(42) = .336$ ,  $p > 0.025$ . Hence, there is no statistically significant difference in mean between the Reader Response and the Philosophy for Children program (P4C) approaches for critical thinking, although the P4C group on its own showed a significant improvement as revealed by the test on hypothesis one.

The results of a paired difference *t* test for the pre and post test for reading skills in the P4C group is shown below. Table 6 shows that there was adequate evidence to reject the null hypothesis,  $H_0$ . This meant that there was a statistically significant mean difference for reading skills in the P4C group,  $t = -3.253$ ,  $df = 22$ ,  $p < 0.025$ .

**Table 6:** The paired difference *t* test for the reading skills of the P4C Group.

	Mean	M.dif f	N	S.d.	t	df	P
Pre TOEFL	35.22	2.39	23	5.27	-3.253	22	.004*
Post TOEFL	37.61		23	4.27			

\* Significant  $\alpha = 0.025$

However, Table 7 indicates that there was no significant difference for the reading skills of the Reader Response group,  $t = -.697$ ,  $df = 19$ ,  $p > 0.025$ . Therefore, the null hypothesis,  $H_0$ , was not rejected.

**Table 7:** The paired difference *t* test of the reading skills of the Reader Response Group.

	Mean	M.dif f	N	S.d.	t	df	P
Pre TOEFL	33.40	0.55	20	5.97	-.697	19	.495
Post TOEFL	33.95		20	4.52			

Hence, it could be further concluded that the P4C intervention was effective since it had enhanced reading skills significantly. The RR intervention had not.

An independent two-tailed *t* test was employed to compare the two groups of P4C and RR for reading skills. Table 8 reveals the result of the *t*-test.

**Table 8:** Two-tailed *t*-test between treatment groups for reading skills.

\* Significant  $\alpha = 0.05$

Treat- ment	N	Mean	M.dif f	S.d.	t	df	P
P4C	23	37.6087	3.228	4.2719	2.352	42	.023
RR	21	34.3810		4.8319			

Table 8 shows the mean score for reading for both groups. As shown, the P4C group scored higher than the RR group. The *t* test also shows that there was sufficient evidence to reject the null hypothesis ( $H_0$ ),  $t(42) = 2.352$ ,  $p < 0.025$ . Therefore, it could be concluded that there was a



statistically significant difference in mean between the Reader Response and the Philosophy for Children program (P4C) in reading skill scores. P4C was shown to be a better approach than RR for enhancing reading skills. This was consistent with the results of tests of hypothesis three.

Table 9 shows the results of the correlation between the TOEFL and the NJTRS scores. The results indicate a statistically significant positive linear relationship between thinking ability and reading ability ( $r = .582$ ). The test showed a significant correlation between thinking ability and reading ability for the P4C group.

**Table 9:** Correlation Table of Post Reasoning Scores (NJTRS) and Post Reading Scores (TOEFL) of the P4C Group.

\*\* Correlation is significant at the 0.01 level (2-tailed).

	NJTRS (Reasoning Score)	TOEFL (Reading Score)
Post NJTRS Pearson Correlation	1.000	.582**
Significant (2-tailed)		.006
N	22	22
Post TOEFL Pearson Correlation	.582**	1.000
N	21	23

The same was true for the Reader Response group, as indicated in Table 9. There was a statistically significant positive linear relationship between these two variables: post- NJTRS scores and post-TOEFL scores. Therefore, it could be concluded that there was a positive correlation between reading ability and thinking ability in both the treatment groups. This confirmed the assumption or premise that language ability is related to thinking ability. Thus, enhancing thinking skills would also improve reading skills and vice-versa. The next section considers findings in the quantitative data.

The qualitative findings from the students' journal entries showed that there were distinct differences in the quantity and type of entries (see Table 10).

In the RR group, the opinions or arguments were related to the story line of the literature, while in the P4C group, the entries were philosophical in nature, for the topics were related to ethical matters or philosophical issues discussed

**Table 10:** General Findings on Journal Entries.

Items	RR	P4C
1. Entries	Short/brief entries About 7 entries per student. Average word count per entry 137. Each entry was based on the story	Longer entries Ranged from 15 – 20 entries per student. Average word count per entry 370. Entries were based on episodes of the text.
2. Style	Majority had narrative, moralizing—class had discussion but not transferred	Tend to have question followed by answer—based on text and pattern of discussion held
3. Outlines	Narratives, what happened in the classroom, moralizing, react to story, characters or lessons held.	Giving opinions, reasoning, discussions, probing and questioning.
4. Discussions/arguments	Lacking (no transfer of the classroom discussion or argument into the journal)	Plentiful discussion, arguments and reasoning. Transfer of discussion or arguments in the classroom was visible in the entries
5. Variety	Lack of questioning, reflection, meta-cognitive, discussion, argument and probing.	Sufficient questioning, reflection, meta-cognitive awareness, and inference of the text, probing, discussion and argument.
6. Moral judgment	Affective response—moralizing	There was affective response—moralizing
7. Improvement	No gradual improvement—in fact some entries superficial and brief.	Gradual improvement was visible especially among students of low proficiency level.

during class or raised in the text. In other words, the text and class discussions determined the topic of journal entries. Further evidence of the influence of the text was that the entries were episodic and had the question format typical of the P4C text.

Another difference was found in the variety among the P4C journal entries. There were examples of reflecting, rationalizing, probing, wondering, and the relating what was discussed with happenings around school. In short, the



P4C entries showed more evidence of critical thinking elements.

The following are some examples of RR entries, presented as they were written:

- Today we learned something about the Necklace. It told us about the story of a husband and a wife who had a different perception. Then after that we had a debate between the girls and boys entitled, "Is Loisel a weak husband?"
- Yesterday, our EST class was a little interesting. We, four Iman, have discussed and learned about a short story called "The Necklace." It a kind of interesting story. It tells us about Matilda who had lost her friend necklace and repay it back with even more expensive necklace.

As shown above, the RR entries did not describe the classroom discussion. This contrasted with the P4C entries, which were about the discussions held during class. The following P4C journal entries clearly show this:

- Then we discussed about, "Do we avoid doing something because of consequences or because it is set by law?" I think both of them are the reasons why we avoid doing something. But people who is more responsible will always think about consequences even though it is not set by law. For example littering rubbish in public places. In some places, there are no signboards which said, "Do not litter" we can do as we like but...
- Next we discussed why in Islam that the deads were buried quickly even though his or her close family who have to arrive late, wants to see his or her face for the last time. After much discussion and arguments, I came to the conclusion that in Islam we have the concept of the soul. A soul is permanent and when the body is dead the soul is trapped until the body is buried. So it is important that Muslims bury the body as soon as possible so that he (the soul) can go to meet his maker, Allah s.w.t.

In the classes we observed—the last of both groups (RR15 and P4CP15)—the lessons held were in the form of Socratic discussions, with the students seated in a circle. Socratic discussion is a form of classroom talk that is moved forward by a series of linked questions. As a qualitative aspect of the study, we analyzed the questions and the events in the classroom in relation to the promotion of critical thinking.

Even though both lessons conducted ostensibly the same sort of discussion, it was obvious that the P4C group was student-led while the RR was teacher-led. In the P4C group, a student was appointed to lead the discussion. He or she informed his classmates of the purpose of the lesson, determined the structure of the lesson, elicited responses from his or her classmates, and nominated a fellow student for a verbal response. The teacher sat behind the students,

and participated at times. In the RR, the Socratic discussion was teacher-led, and the teacher addressed the whole class. He or she determined the flow of discussion, elicited questions, and did most of the probing for further meaning. However, for both groups, the teacher concluded the lesson or topics discussed.

Since the P4C class was student-led, the onus was more on the students—especially the moderator—than on the teacher to get the discussion going. This contrasted with the RR class, which remained teacher-centered. In the P4C group, students felt more responsibility to bring about interaction among themselves than between students and the teacher. This was not so in RR, where interaction among students was not emphasized, for the lesson or method did not give much opportunity to bring this about. In the P4C lesson observed, there were two distinct forms of interaction which took place during the discussion—one among students, and the other between the students and the teacher. Both interactions were prominent features of this P4C class. In the RR group, interaction among students was not frequent, and what exchanges there were, were short and few. Secondly, there was a distinctly higher quality of cognitive interaction among students in the P4C group than in the RR.

In the P4C classroom, pupils were involved in eliciting questions from each other and determining the topic of discussion, after gleaning questions from the text that they had read. In these arguments for what to discuss, elements of critical thinking were obvious. Students were making judgments, drawing from personal experience, linking ideas and facts, and evaluating the contribution of their peers. In other words, the students manifested reasoning skills, made critical thinking responses and assessments, and enquired collaboratively. Community of inquiry behaviors were evident in eliciting opinions or posing questions for reflection, countering opinions with arguments, or disagreeing with the view points raised by their peers with relevant statements, and justifying judgments made. In the RR groups on the other hand, counter-argument emerged most typically in differences of opinion between the two subgroups of boys and girls. In all cases, the exchanges were brief and consisted of short statements or incomplete sentences that were not clear. The discussion tended to lack seriousness and depth, each group arguing in adversarial mode, without a concern to justify or even to give reasons.

The findings here showed that in the P4C groups critical thinking was enhanced when the students themselves raised questions and thus became a community of inquiry. Some of the characteristic of the nascent community of inquiry observed here were the ability of its members to make relevant statements and to probe with questions, to elicit questions or topics for discussion in the form of questions, to draw on personal experience, to use previous knowledge as evidence, to clarify ideas, to link facts or ideas, to justify examples, and lastly to make judgments and critical assessment.

Another contrasting pattern was the number of



elicitation acts (a question or statement which functions to gain a verbal response from another speaker) in the pupils' responses. The P4C pupils asked more questions of their fellow classmates than did the RR. The P4C group asked 14 questions, while the RR group asked only 5.

The type of questions raised also differed. Table 11 and Table 12 itemized the questions raised by the respective groups. As shown in Table 11, the P4C group showed a variety of questions posed compared to the RR group. The questions raised by the P4C group were higher-level, cognitive (divergent) questions, unlike those raised in the RR.

**Table 11:** Questions raised by pupils during P4C class 15.

1. What can we get from this passage?
2. Why do people feel ashamed when they think about their past?
3. Do we feel ashamed when we think of the past?
4. Do we avoid doing something because of the law or the consequences?
5. Can someone explain what this question (above) means?
6. Are we afraid of the law or are we afraid of the consequences?
7. Can some one describe some of the things or events that happened in the past that embarrassed you?
8. How do the girls study to make it memorable?
9. How? (to make studying interesting). Give examples
10. Do you think it is good to have such a law? (referring to a ban on chewing gum in Singapore)
11. So is the barbecue... so why don't we have a law against barbecue?
12. We have a right even to barbecue isn't it?
13. Will it affect the ozone layer? Will smoke really affect the ozone layer?
14. Are you sure the smoke, the barbecue smoke doesn't have CFC?

Qualitative data from teachers observation notes and journal entries noted a marked distinction between the two groups in the structure of the lesson. The structure of the P4C lessons was ritualized—that is, it had a set of mandatory routine procedures involving three stages. The first stage, reading, was followed by the next stage, elicitation of questions. The last stage was group discussion. In contrast, the structure of the RR lessons was mixed, for there was variety in its lesson plans. There were role-plays, interviews, a debate, a Socratic discussion and a trip to the theater—all methods suggested by proponents of the RR approach, and aimed to enhance students' response towards literature. Although the underlying philosophy of the RR approach is to enhance students' response to a literature text

**Table 12:** Questions raised by pupils during RR class 15.

1. (If) the man can marry four, why can't the women marry four?
2. Why don't you stay married with one? Why marry another (the other one)?
3. She doesn't want to be shared. So how can you be fair?
4. Why you marry her in the first place?
5. How to be fair with one wife? Teach me how to be fair.

through their making their own interpretations, the critical thinking element in such an approach comes through the teacher's probing questions during the response stage, or through discussions throughout the lessons.

The activities carried out succeeded in eliciting student responses, although the degree of response differed. However, despite the good response at the initial stage (especially during role play), the level of response could not be sustained in the subsequent discussion. Questions from the teacher or subsequent probing after a role-play, for instance, were not keenly answered or even responded to at all. At times, the teacher ended offering probing question after question, sometimes with no takers. This was significant, for as explained earlier, the input of the critical thinking element in Reader Response is expected to come from the teacher's probing questions during the response stage and subsequent discussions or lessons.

The underlying principle of P4C is to enhance the critical thinking of its students through two approaches: leading students to ask their own questions, and thereby create a community of inquiry. This was carried out in this case through the use of the text *Lisa*, combined with the teacher assuming the role of one of the community members, as well as acting as a guide in the classroom discussion, and helping to shape the formation of a community of inquiry. For this, the role of the teacher in offering probing questions during the discussion was essential. In fact, in both approaches the teacher was seen to constantly probe the students; the number of probing questions in each group was almost equal. However, constant probing by the teacher was not a guarantee that students would acquire the ability to probe. In the RR group, student probing was not as frequent as in the P4C group. In the P4C group, the students probed increasingly throughout the testing period, and were seen to probe their classmates especially in discussions when the student became the moderator. This occurred at a later stage of the study, and corresponded to the growth and maturation of the community of inquiry.

In general, depth and a seriousness of discussion with classmates was observed more frequently in the P4C than in the RR groups. The topics discussed and the lack of dramatization in the P4C lessons led to the more serious nature of the discussions, while in the RR groups, role-playing either by the teacher or the pupils enlivened the mood of the class, which was an important trigger for pupils' responses, but



I liked the question.

contributed to less depth of discussion. The P4C pupils were also observed to elicit more questions from their peers during class discussion. Elicitation of questions occurred during the questioning and discussion stage through either teacher-pupil or pupil-pupil interaction. Elicitation of questions from and between pupils was of a lesser degree in the RR group. Again more questions from pupils were raised during specific types of lessons such as the debate, the role-play lessons, the interviews, and to a lesser extent, the Socratic discussion.

The questions raised in each group were highly dependent on the text. This was obvious in the P4C group, where students gleaned questions from the text or backtracked to the text in order to explain something. As such, the content matter and flow of discussion in the P4C group were highly dependent on the text and its various interpretation. As we have seen, the P4C text contained topics pertaining to ethical issues, but the discussion of issues in the classroom depended on selection and interpretation. In fact, the issues raised in the P4C discussions cut across all disciplines—including morality, civics, science and technology, mathematics, social analysis, personal development and the environment. In the RR group, students were seldom seen to refer or backtrack to the text for any questions, which came mostly from the teacher. Even then, the questions raised were related to the story line of the piece of literature discussed. Thus, the type of discussion was dependent on the approach.

The types of questions raised in each group differed as well. Those raised in the P4C group were more diverse. Some of the topics were related to moral rights and the sense of justice, and this influenced the type of discussion that followed. The questions raised in the RR group focused around the story line of the literary piece. They also included questions about events, characters or the setting of the story. It was also observed that the RR students rarely had difficulty in answering the questions raised, whereas in the P4C lessons, the students took time to reflect on questions, posed by themselves, for which there was no clear answer. The higher cognitive level of the questions raised in the P4C group is evident from an examination of those listed in Tables 11 and 12. Nor was there any doubt that the types of questions raised affected the nature of the discussions that followed.

As is clear from the literature review undertaken for this study, the community of inquiry is essential to the P4C

methodology in fostering critical thinking or awareness. However, in this study, evidence of the formation of community of inquiry behaviors and dispositions was only detected at the later stage—particularly in fourteenth, fifteenth and sixteenth lessons. In fact, three stages of development were observed over the course of the P4C lessons. The initial stage lasted from the first through the twelfth lessons, a transition state was evident in the thirteenth lesson, and the third stage—the nascent stage of the community of inquiry—emerged in the fourteenth through sixteenth lessons.

During the initial stage, the discussion was teacher-led, and the latter played a central role in guiding the discussion. The students were relatively passive, participation was lackluster, and there was a tendency for chorus response from pupils, or a one-word response after being probed individually. Often the responses came from the same persons, typically the average or high-achievers. The low achievers responded curtly (one word response) when the teacher deliberately probed them. Thus, up until the twelfth lesson, no community of inquiry was detected. Because of this, the teacher decided to appoint a student as a moderator during the thirteenth lesson, and subsequently the students' participation increased as the discussions became student-led. However, the teacher did not participate in these discussions and all, and as a result, there was some confusion, and the students' arguments tended to be ambiguous and vague. From the fourteenth to the sixteenth lesson, the teacher became involved again, and it was at this point that signs of community of inquiry were detected. Negotiation of meaning was observed between students, moderator and teacher during the discussion. Because of its late formation, we have characterized this form of community of inquiry formed as nascent. During this stage two types of interaction were observed—student-student and students-teacher. It was also observed that students began to relate what was being discussed at the moment to previous topics. In the previous lessons, only the teacher had been doing this. In this stage, the responsibility for initiating and maintaining the discussion going was on both the students *and* the moderator.

While community of inquiry behaviors and dispositions were observed in the last sessions of the P4C group, none were observed in the RR group. The discussion was teacher-led throughout, and student participation during discussion was generally lacking. As stated earlier, students responded well during role-play, but their enthusiastic re-



sponse could not be sustained during the discussion stage. No student-student interaction was observed, and the responsibility for moving the discussion forward was clearly considered to be the teacher's. Similar observations were made during the initial stage of P4C lessons.

### Conclusion

There is no doubt that, comparatively speaking, the P4C methodology enhanced students' critical thinking more than the RR. The critical contributory factors were the quality of discussion, the emphasis on developing one's own questions, and the character of the text. The latter was a story in which were embedded philosophical issues, and the characters of the story were about children who were themselves thinking critically, and thus represented role models for the students. The philosophical issues embedded in the text were often picked up by the students and formed the raw material for their questions, which then became the source for in-depth discussion. In other words, the text guided the class in what to discuss and how, and the elicitation of questions was the preparatory stage or impetus for subsequent discussion. This explains the result of the *t* test—a statistically significant difference in the pre and post test skills in the P4C group but not in the RR group. That is, the P4C intervention was effective in significantly enhancing critical thinking skills, while the RR was not so effective.

However, a comparison between the two approaches shows that improvement in critical thinking skills was not significant in either. The qualitative findings suggest that community of inquiry understood as a set of dispositions and behaviors was still in the nascent stage in the P4C group, suggesting a reason for the absence of significant improvement in critical reasoning skills in the latter. This is understandable, given the importance which the literature attributes to the pedagogy and group process of community of inquiry in promoting these skills.

### References

- Abott, J. 1997. Critical transitions: human capacity development across the lifespan. <http://www.21learn.org/>
- Benton, M. 1996. Reader-Response criticism. In Hunt, P. *International companion encyclopedia of children's Literature*. London: Routledge.
- Gelberg, D.S. 1993. Education reform: then and now- the influence of business in education in the 20<sup>th</sup> century America. Ph.D. Dissertation. University of Boston. Proquest Digital Dissertation International Abstracts. AAT 9401888. <http://wwwlib.umi.dissertations/gateway>.
- IAPC/ Program & history: about the program. 2002 <http://www.montclair.edu/pages/iapc/history.html>
- Langer, J.A. 1997. *Literacy acquisition through literature*. New York: The National Research Center English language and Acquisition. <http://cela.albany.edu/acq/index.html>
- Lim, T. 1998. Ascertaining critical thinking and formal reasoning skills of students. *Research in Education*. 59 (1): 9.
- Lipman, M. 1983. *Lisa*. Upper Montclair, N.J: The IAPC.
- Lipman, M. 1991. *Thinking in education*. Cambridge: Cambridge University Press.
- Lipman, M. 1993. *Thinking children and education*. Iowa: Kendall/Hunt Publishing Company.
- Malaysian Ministry of Education. 2000. *Selected poems and short stories: Form 4 Literature in English for upper secondary schools*. Kuala Lumpur: Dewan Bahasa dan Pustaka.
- Miller, H. 2002. Reader Response in secondary and college Classroom. *Journal of Adolescent & Adult Literacy*. 45. (4): 327 – 329.
- Nussbaum, E. 2002. The process of becoming a participant in small-group critical discussions: A case study. *Journal of Adolescent & Adult Literacy*. 45 (6): 488.
- Page, M.L. 2001. Weaving a teaching tapestry: the dialogic act of teaching multicultural literature in a middle school. Ph.D. Dissertation. University of Wisconsin. Dissertation Abstracts International. <http://wwwlib.umi.dissertations/gateway>.
- Reed, R.F. 1992. On the art and craft of dialogue. In Sharp, A.M. & Reed, R.F. (eds). *Studies in philosophy of classroom*. Philadelphia: Temple University Press.
- Resnick, L.B. 1989. Overview toward the thinking curriculum. In *Toward the thinking curriculum*. Pittsburgh: ASCD Pub.
- Rosnani Hashim & Suhailah Hussein. 2003. The teaching of thinking in Malaysia. K.L: Research Centre International Islamic University.
- Shipman, Virginia. 1983. *New Jersey Reasoning Skills Test*. N.J: The IAPC.
- Suhailah, H. 2001. Teachers' perceptions of the teaching of thinking and their practices in secondary schools in Klang, Selangor. Unpublished Masters in Education ertation. Gombak: International Islamic University Malaysia.

peopoll canq ther mind.