

Philosophy for Children (P4C) in Improving Critical Thinking in a Secondary Moral Education Class

Hafizhah Zulkifli

Universiti Kebangsaan Malaysia
Bangi, Malaysia

Rosnani Hashim

International Islamic University Malaysia
Gombak, Malaysia

Abstract. The purpose of this study was to identify the effect of Philosophy for Children (P4C) in improving critical thinking among students. The study was conducted via the quasi-experimental research comprising of 27 students placed in the experimental group while the remaining 34 students in the control group. The instrument used was the *Ujian Kemahiran Menaakul* (Test for Reasoning Skills) Centre for Teaching Thinking (UKMCTT). Based on the statistical t-test on critical thinking, the treatment group scored a higher mean in the post-test mean score compared to the control group, which suggests that P4C had helped to improve students' critical thinking. The application of P4C allows students to think or reflect on the consequences of the action or assumption they made in their discussion. Simultaneously, it provides career advancement for the teachers by providing the classroom with a new strategy and renewed energy and enthusiasm.

Keywords: Philosophy for Children (P4C); Critical Thinking; Moral Education; Students; Test for Reasoning Skills

1. Introduction

Teaching student's critical thinking will strengthen their logical skills and ability to reason (Lipman, Sharp, & Oscanyan, 1980). Effective teaching is comprised of three strategies, i.e., dialogue strategies, stimulating critical thinking, and encouraging discussion in the classroom (Hashim, 2013). Moral and academic excellence are profoundly affected by effective teaching (Solomon, Watson, & Battistich, 2001), and teachers who have positive attitudes about their students can enhance the students' achievement and ethical behaviour (Haberman, 1999). Teachers need to have content knowledge as well as to associate caring classrooms, achievement, and prosocial character. Pedagogical skills are also

necessary, and they are expected to be committed to providing caring climates as a teaching practice (Che Abd Rahman, 2007).

In Malaysia, thinking skills are a significant agenda in national education and are eminent (Yen & Halili, 2015). Malaysia has put a significant effort in enhancing thinking skills with the implementation of structural reforms by the Ministry of Education (MOE) through the Integrated Curriculum for Secondary Schools (KBSM). The concept of critical thinking was incorporated in some major national plans including in KBSM in 1988, Vision 2020 in 1991, the Critical and Creative Thinking Skills (KBKK) in 1996, and the concept of “smart school” to produce human capital with high-level thinking ability introduced in 1997 (Ministry of Education, 2012).

In 2012, the Preliminary Report of the Malaysian Education Blueprint 2013–2025 was introduced by MOE that emphasised on six critical attributes needed by the students to thrive in the globally competitive market such as knowledge, thinking skills, leadership skills, bilingual proficiency, ethics and spirituality, and national identity (Yen & Halili, 2015). Nevertheless, in reality, the majority of teachers have failed to engage students in effective teaching as they still employ the lecture method instead of promoting the culture of Higher Order Thinking Skills (HOTS) (Ivie, 1998; Ministry of Education, 2012; Zohar & Schwartz, 2005).

Moreover, the majority of teaching methods involved memorising (Hoon, 2010; Glickman, 1991), taking notes (Balakrishnan, 2002, 2009), possess lack of activities (Appoo, 2009), emphasises on exam (Barone, 2004), asked lower-order thinking questions, and usually used the lecture technique (Che Abd Rahman, 2007). This pedagogy assumes teachers as a source of knowledge, with student engagement at a minimal level and indirect interactions between students.

In addition, Nachiappan, Sinnasamy & Suffian (2017) contend that the problems that teachers have to overcome were the inability to relate to the moral values in real-life situation among students, lack of teaching aids, lack of mastery in lesson delivery among teachers, lack of interest from students, inability to explain about moral issues among students, problem in mastery of language among students, presence of teachers with no Moral Education field background and the inability to understand and master the moral values among students. Moreover, the influence of students’ mother tongue also has become an alarming issue among teachers. Students also tend to have an attitude towards Moral Education which shows their lack of concern which leads to them being unable to pay attention. Then, lack of optional teachers in Moral Education, lack of references and uncondusive classroom surrounding can cause problems in Moral Education.

In light of such issues, we can conclude that students have not achieved a higher level of critical thinking. Thus, students need to be catered with a

pedagogy that provides them with the opportunity to enhance their critical thinking. So far, not many researchers have paid attention to the usage of philosophical inquiry such as P4C as an alternative methods in Moral Education in order to enhance critical thinking. Based on these stated gaps, thus, this study aims to examine the effect of P4C on the students' critical thinking level after the P4C is implemented in the classroom.

2. Philosophy for Children (P4C)

Several critical thinking programmes are available to teach students to think critically. The 'Teaching for thinking' approach was developed by teachers across the world in different ways. They integrate the approach into everyday teaching to create 'thinking classroom' and develop whole school policies to create 'thinking school'. Lateral Thinking, which was promoted by Edward De Bono (Bono, 1995) is one of the examples that has inspired a wide range of curriculum and programme developments, while the Instrumental Enrichment Program by Reuven Feuerstein (Feuerstein, Falik & Feuerstein, 2015) is a programme for struggling learners. It also includes 'cognitive acceleration' approaches (such as Adey and Shayer's Cognitive Acceleration through Science Education), which are called 'brain-based' approaches like the 'Accelerated Learning' and 'philosophical' approaches, such as Lipman's Philosophy for Children, which unlike other approaches, not only aims to develop the intellectual aspects of thinking, caring and collaboration, as well as critical and creative thinking, but also the moral, social, and emotional aspects (Fisher, 2013). Philosophy for Children (P4C) was created in the late 1960s by Matthew Lipman and his colleagues as a coherent programme in teaching thinking at the Montclair State University, United States. This programme was created due to the realisation by Lipman (1993) who was a professor teaching philosophy at Colombia University at the time that the undergraduates lacked in reasoning and judgment (Naji, 2005).

According to Lipman et al. (1980), the P4C was built upon the recommendation of John Dewey (1938) and the Russian educator, Lev Vygotsky (Minick, 2005). They emphasised on the necessity of teaching for thinking and not for merely memorising (Lipman, 2017). He mentioned other philosophers and psychologists who instigate influence in P4C such as Justus Buchler, Jean Piaget, Gilbert Ryle, George Herbert Mead, and Ludwig Wittgenstein (Lipman, 2017). It is essential to highlight that the aim of P4C "is not to make children into little philosophers but to help them to think better than they now think. The more quickly they can adapt to philosophy, it will emphasis mental acts, thinking skills, reasoning and judgment" (Lipman, 2017, p. 9).

In providing and pursuing a P4C curriculum, the Institute for the Advancement of Philosophy for Children (IAPC) began to implement the P4C programme in kindergarten to the twelfth-grade classroom in Montclair. For the first draft, he published a novel for children entitled *Harry Stottlemeier's Discovery* to foster children's critical thinking, discussion, and questioning (McLeod, 2010). More philosophical children's novels were published in the following years by Lipman such as Suki, (1978), Mark (1979), Pixie, (1981), Lisa (1983), Elfie (1987);

Kio & Gus, 1982). Every novel was different as each novel was written to suit different age groups and schooling grade, as shown in Table 1.

Table 1. Philosophical Novel

Age	Children's novel	Teacher's manual	Philosophical area	Educational area
3-6 years	Doll Hospital	Making sense of my word	Concept formation	Basic concepts- what is true, good, real, beautiful?
6-7 years	Elfie	Getting our thought together	Reasoning about thinking	Exploring experience
7-8 years	Kio and Gus	Wondering at the world	Reasoning about nature	Environmental education
8-10 years	Pixie	Looking for Meaning	Reasoning about language	Language and arts
10-12 years	Harry Stottlemeir's Discovery	Philosophical Inquiry	Basic reasoning skills	Thinking and logic
12-13 years	Lisa	Ethical inquiry	Reasoning in ethics	Moral education
14-15 years	Suki	Writing: How and Why	Reasoning in language	Writing and literature
16 + years	Mark	Social inquiry	Reasoning in social foundation	Social studies

Source: Fisher (2013)

Each novel, as explained by Lipman (1988) revolves around a common theme, which is the workings of the human mind, and the primary purpose of the theme is to serve as springboards for intellectual debate. However, one of the drawbacks of the novels was that they are not good stories in a literary sense, as they do not hold the interest as stories. Nevertheless, Lipman saw this as an advantage. Most children's book and stories do not contain a rich range of philosophical questions, and they do not provide models for children as enquiring thinkers. He believed that children's existing literary experience lacked in intellectual stimulation, causing the link between reading and thinking to be disconnected. Children usually do not think about what the words mean for the story, and they eventually believe that reading is only about following words in a book (Fisher, 2013).

Besides, Lipman's 'philosophic novels' (1988) by contrast, are seeded with puzzles, questions, and problems of meaning. There is an educational purpose to stimulate questioning and philosophical discussion. They present models of reasonable and thoughtful discussion between children that Lipman hoped his

readers would emulate in their discussion of questions arising from the text (Fisher, 2013).

Furthermore, a philosophical novel is more effective than a philosophical context because it is far more enjoyable to be conveyed through stories. Children will be more motivated to think and inquire if the stories focus on issues and events that are related to their personal experience but at the same time, are intriguing and contestable. The philosophical text is an attempt to present philosophy logically and comprehensively despite the lack of experience. Stories will be used as a springboard or trigger for their further inquiry by children who are in the process of building their communities of philosophical inquiry (Sharp, 2017).

3. Methodology

Research Design

A quasi-experimental design was chosen for the assessment of this study. The researcher used existing classroom to avoid disrupting classroom learning (Creswell, 2005) and it is usually not possible for educational researchers to randomly assign pupils to groups or to design a study in which the control group is taught the irregular programme from the prescribed curriculum (Ann, 1993).

In the quasi-experimental design, the participants were divided into two groups with group A1 as the experimental group and group B1 as the selected group without random assignment. A pre-test and post-test were applied to both groups, and only the experimental group will receive the treatment (Creswell, 2009). In the treatment group of this study, students were taught through the use of P4C while in the control group, students were offered the regular formal Moral Education course. Both experimental and control groups consisted of form two students.

Group A1	O ₁	X		O ₂
Group B1	O ₁			O ₂

Group A1: Experimental Group (P4C)

Group B1: Control Group (conventional learning)

O₁: Pre-test

O₂: Post-test

X: Treatment

Sample

The population of this study was form two Moral Education students. The samples of this study comprised of two classes of Moral Education students in one secondary public school in Gombak. The sample was chosen due to the diversity of non-Muslim students in the school, the study's ability to attain

cooperation from the school administrator, and the conducive environment for teaching and learning offered in the school.

The form two students were selected because they were not involved in any major compulsory examination. The students were selected from schools that consist of a minimum of ten non-Malay or non-Muslim students as Moral Education is only offered to non-Muslim students. According to Berg and Latin (1994), there should be at least ten subjects per group when researching group comparison.

The samples from the two classes of form two consisted of 61 students. The total number of students for the experimental group was 27, and the control group consisted of 34 students. The sample experimental group consisted of 13 males and 14 females, which consisted of 19 Indian students and 8 Chinese students. Meanwhile, the control group consisted of 13 males and 21 females with 30 Indian students and 4 Chinese students.

The samples were not randomly assigned to the two groups as this study cannot artificially create groups for the experiment to avoid the disruption of classroom learning. The form two classes were in the afternoon session. The two classes of Moral Education involved were 2A and 2C, where 2A stands for *Aktif, Amal, and Amanah* that comprised of excellent and moderate students based on their form one year-end exam achievement. Meanwhile, 2C that represent *Ceria and Cerdas* consisted of moderate and low achievement students based on the same assessment. 2A was assigned as the control group and 2C as the experimental group.

The Moral Education class provides students with multicultural ethnics, different religious beliefs, different spoken languages, and dialects. In addition, the Moral Education class emphasises on Moral Reasoning, Moral Feeling and Moral Action that is suitable for this study.

Instrument

The *Ujian Kemahiran Menaakul* Centre for Teaching Thinking (UKMCTT) used to examine the level of critical thinking among students is an adaptation from the New Jersey Test of Reasoning Skills (NJTRS) that was developed by Dr. Virginia Shipman, a senior research psychologist in the Education Department in New Jersey (Shipman, 1983). It is a test of the ability to reason, rather than a test of scientific inquiry or judgment. It is clear with a written Flesch Reading level of 4.5, and reliability ranging from 0.84 to 0.91 and compares favourably with other thinking tests such as the Cornell Critical Thinking Test and the Whimbey Analytical Skills programme.

The test is recommended for high school and colleague students. It has been widely used with groups of foreign language learners in high schools in many parts of the world that conducted the P4C programme in the past 30 years (Institute for the Advancement of Philosophy for Children). These countries include Singapore (Lim, 1994) and Malaysia (Hashim, Hussien, & Imran, 2014).

A reliability test was conducted for this study and its Cronbach alpha for internal reliability was 0.73.

The UKMCTT had been translated and had 33 multiple-choice items. Some of the reasoning skills included are reasoning with the relationship, identifying good reasons, inductive reasoning, recognizing improper questions and analogical reasoning, and syllogistic reasoning. The duration of the test was one hour.

Data Collection Procedures

Initially, the researcher met with the principal to discuss the class and timetable that would be suitable for the study. Next, the researcher entered the class and built rapport among the students in the classroom for about a month. Accordingly, the researcher gave critical thinking instruments for the pre-test while the teacher started implementing the intervention module for 11 sessions. After the teacher completed the 11 sessions of intervention, the researcher again gave out the critical thinking instruments to the students in the post-test.

Data Analysis Procedures

The researcher used inferential statistic comprised of independent-samples t-test and paired sample t-test on UKMCTT instruments. As stated by Pallant (2007), researchers use the independent sample t-test when they want to compare the mean scores of two different groups of people or conditions. On the other hand, when the researchers want to compare the mean scores for the same group of people on two different occasions, or when there are matched pairs, a paired sample t-test is used. In this research, SPSS was used to analyse the data.

Validity

Validity in quantitative data, as mentioned by Creswell (2005) refers to when individual scores from the instrument are meaningful and make sense. The validity of an instrument also enables the researcher to draw a reasonable conclusion from the sample of the population. The quasi-experimental approach introduces more threats to internal validity than the actual experiment. This happens because the researcher does not randomly assign participants to groups. A threat to validity refers to when design issues can threaten the experiment conclusions reached from data, leading to a false reading and plausible cause and effect between the treatment and the outcome (Creswell, 2005).

The threats to internal validity threatened the researcher's ability to draw the correct cause and effect of inferences that arise. This is profoundly affected by the experimental procedures or the experiences of participants. History, maturation, regression, selection, mortality, and interaction with selection are among the potential threats to internal validity (Creswell, 2005; Campbell & Stanley, 1963; Christensen, 2004).

Table 2 shows a list of internal validity threats and suggested means of control. First, history refers to any situation or event that coincides with the treatment

given to the group, which could have a similar effect. In this experiment, historical threats to internal validity were (i) students who participated in the P4C programme were also eligible for and probably participated in the academic excellence programme, which also aims in improving higher-order thinking skills, particularly during school holidays and at times, on Fridays; and (ii) the unequal syllabus for the control group and the treatment group. This is because the schedule for the experimental group falls typically on the same day with a public holiday, and as a result, the class would miss or skip one topic. Schedule reshuffling was thus done to overcome this problem.

The second threat, maturation, refers to any changes that happen through the natural process in everyday lives to both the experimental and control groups. For this experiment, having the experimental and control groups would help to minimise this problem because if such changes happen to members of the treatment group, it might affect the members of the control group as well. Christensen (2004, p. 2002) asserts that “In order to determine the effect of a programme, a control group that did not receive the treatment would also have to be included to control the potential rival influence of maturation”.

The third is the instrumentation that refers to changes that happen in the measurement procedure. The validation and reliability procedure of the instrument would help minimise this problem. Next, for the biosocial attributes, the researcher observed differences in terms of race and religion. To control this threat, the researcher must be open-minded and tolerant. The researcher also looked into the students’ background and understood the diversity of their culture.

The fourth threat, attrition or mortality, refers to any experimental subjects that drop out of the study before the experiment has been completed. To avoid this issue, the students were reminded that their attendance was compulsory. Nevertheless, some of the students did not show up during the experiment as they were involved in a chess tournament and running events, for instance. There was also an instance in which only Chinese students attended the class since there was an Indian festival that was not declared as a public holiday.

The fifth is an imitation of treatment. Members of the treatment group could communicate with each other and may pass on the experimental stimulus of one group to the other. This might influence the behaviour of the subjects. Again, administrative control and close supervision were essential to avoid any obstruction.

Table 2. Threats to Internal Validity

Elements	Threats	Suggestion
History	Students involvement in an academic excellent programme Unequal syllabus	The participation in the excellent programme to be minimized Schedule to be reshuffled

Maturation	Any changes that happen through the natural process in everyday lives to both the experimental and control groups	The use of the experimental and control group
Instrumentation	Any changes that happen in the measurement procedure	Validation and reliability procedure of the instrument
Mortality	Drop out of the study	Attendance of participants to be made compulsory
Imitation of treatment	Members of the treatment group could communicate with each other	Administrative and close supervision

In addition, the threats to external validity are problems that threaten the ability to draw correct inferences from the sample data to other persons, settings, and past and future situations such as the interaction of selection treatment, the interaction of setting and treatment and interaction of history and treatment (Creswell, 2005). In terms of external validity, the findings from this research are limited in "generalisability"; thus, it cannot be inferred to other Moral Education students in Malaysia. While this study and previous findings indicate a pragmatic superiority of the moral reasoning learning module, it does not suggest that competition and individualisation should be overlooked or ignored.

Pilot Test of Critical Thinking Instrument (UKMCTT)

In this critical thinking instrument, the researcher used test and retest reliability. For the UKMCTT, the pre-test was given in June and the post-test in September. The pilot test involved 27 students, and the results showed that the Pearson correlation was .47, $p < .05$. It shows that the UKMCTT instrument is reliable and suitable to measure critical thinking among students.

4. Findings

The students were tested for critical thinking using the IIUM Centre for Teaching Thinking Test of Reasoning Skills (*Ujian Kemahiran Menaakul* Centre for Teaching Thinking, or UKMCTT) during the pre-test and post-test. The students were required to answer within an hour. The students' scores were calculated and converted to percentages.

Test of Effect of Treatment on Critical Thinking

A paired sample t-test was used to compare the critical thinking score among each member of the treatment group before and after the programme. The null hypothesis is:

Ho1 There is no statistically significant difference in the means of the pre-test and post-test for each of the treatment groups in the test of critical thinking.

Table 3 shows the results of the paired sample t-test of the pre-test and post-test mean scores on critical thinking for the treatment group. There was a statistically significant difference in the mean score, $t(26) = -4.54$, $p < .05$. The null hypothesis failed to accept H_0 . This implies that the P4C had helped to improve students' critical thinking in the treatment group.

Table 3. Paired Sample T-Test between Pre-Test and Post-Test Mean Scores in Critical Thinking for Treatment Group

N=27		Mean	MD	SD	SE Mean	t	df	Sig
Pair 1	Pre-test	21.0	3.2	3.58	.68	-4.54	26	.000
	Post-test	24.2		1.99	.38			

Independent Sample T-Test of Pre-Test Score on Critical Thinking between Treatment Group and Control Group

An independent sample t-test was performed to test for equivalence between experimental and control groups. The hypothesis for the independent sample is:

H_{02} There is no statistically significant difference in the level of critical thinking between the treatment group and the control group based on the pre-test score.

The results (see Table 4) show that there was no significant difference in mean scores for the treatment group ($M = 21$, $SD = 3.58$) and the control group ($M = 19.1$, $SD = 4.06$); $t(59) = 1.83$, $p = 0.072$ (two-tailed). The magnitude of difference in the means was very small ($\eta^2 = .009$). The null hypothesis, therefore, fails to reject, and there is no significant difference in the pre-test scores between subjects in the control and the experimental groups. Therefore, the results show that the subjects in both the control and the experimental group were equivalent before treatment.

Table 4. Independent sample t-test of Pre-test Mean Score on Critical Thinking between Treatment Group and Control Group

Group	N	Mean	SD	SE Mean	t	df	Sig (2 tailed)
Treatment	27	21.0	3.58	.68	1.83	59	.072
Control	34	19.1	4.06	.69			

Independent Sample T-Test of Post-Test Score on Critical Thinking between Treatment Group and Control Group

An independent sample t-test for the post-test mean scores of the treatment and control groups was conducted to test whether the intervention of the Pedagogy was significant.

Ho3 There was no statistically significant difference in the level of critical thinking post-test mean scores between the experimental group and the control group.

The results in Table 5 illustrated that there is no significant differences in scores for the experimental group ($M = 24.2$, $SD = 1.99$) and control group ($M = 22.05$, $SD = 2.92$); $t(59) = 3.34$, $p = .001$ (two-tailed). P-value was smaller than .005. The magnitude of difference in the means (mean difference = 2.20, 95% CI; 0.65 to 0.88) were small (eta squared = 0.01). The p-value was below 0.05. Therefore, the hypothesis can be rejected. It can be concluded that there is a significant difference in the post-test mean scores for the level of critical thinking between the treatment and control group on the UKMCTT test. It implies that the P4C approach is effective in helping students to improve critical thinking.

Table 5 Independent sample t-test of Post-test Score on Critical Thinking between Treatment Group and Control Group

Group	N	Mean	SD	SE Mean	t	df	Sig (2 tailed)
Treatment	27	24.25	1.99	.38	3.34	59	.001
Control	34	22.05	2.92	.50			

5. Discussion

The result of critical thinking *t*-tests showed that there was a statistically significant difference in both pre-test and post-test of the critical thinking mean scores of each member of the treatment group. This implies that P4C has improved the students' critical thinking. The preparation of a lesson plan before teaching will also help to improve students' critical thinking. The teacher needs to imagine what the students will try to ask and expect during the preparation of the lesson plan that includes several critical thinking stimuli, questions, discussions, and exercises.

These results are supported by Sasseville (1994), who researched logical skills using NJTRS Instrument and Piers-Harris (1969) for self-esteem instruments on five experimental groups and four control groups. The results indicated that there was a significant difference with increments of 6.28%, 8.66%, and 7.22% for three out of five experimental groups on logical skills while increments of 7.42% and 6.6% on two treatment groups for self-esteem instrument.

Similarly, Othman (2005) and Othman and Hashim (2006) claimed that there was a statistically significant difference of mean for the reasoning skills in the P4C group using NJTRS ($t=.993$, $df=21$, $p<.025$) using a paired sample *t*-test, suggesting that P4C has affected the change in students' critical thinking. The findings also coincide with the findings by Hashim (2002), who applied a quasi-experimental design study among year five students in a private Islamic primary school and discovered that in the pre-test and post-test using NJTRS,

there was a small statistical increase in the mean difference ($t=-2.89$, $df=29$, $p<.009$). P4C also showed a significant difference in five of nine subjects for the upper secondary school. There were significant differences in the pre-test and post-test for History, Oral English for the best class and weak class, oral Malay, and English comprehension. The scores for Science increased but not significantly (Hashim, 2017; Hashim et al., 2017).

Moreover, there was a noticeable improvement for the primary school in all three subjects, which were English composition, Malay composition, Malay comprehension, and Islamic Studies. However, there were no significant differences in all subjects for the lower secondary students in both pre-test and post-test scores, except for Science in 2G, which was a weaker class in comparison to the other two classes, which were the two best classes in their level. Based on the test scores, it can be inferred that P4C has a great potential in increasing not only the soft skills but also academic achievement in the long term (Hashim, 2017; Hashim et al., 2017).

Lam (2012) also supports that P4C can stimulate a student's ability in critical thinking from a large variety and quantity of cognitive behaviours characteristic. Other than that, it could also be concluded that P4C plays a huge role in order to develop a student's critical thinking.

In addition, experimental research by Marashi (2008) conducted with 30 students in the experimental group and 30 students in the control group claimed that students in the experimental group showed a significant improvement in their reasoning skills using the New Jersey Test of Reasoning Skills (NJTRS) instrument.

Research by Ann (1993) revealed that P4C could enhance critical thinking among students. Ann examined the impact of a P4C novel, *Harry Stottlemeier's Discovery* (1976), on the critical thinking of fourth and fifth-grade gifted students in over 12 weeks. She used the NJTRS instrument as a standard test to measure the improvement of critical thinking.

The research by Allen (1988) found a significant improvement in logical reasoning in the treatment group consisting of 23 students. This study coincided with a study by Burnes (1981) who reported that P4C showed a significant result in reasoning skills among grade 5 and 8 students in Minnesota for two years, then in the third year, the results showed an increase in reasoning and also reading comprehension. Based on a few past research, most of them conclude that there are improvements in critical thinking from the use of inquiry in P4C (Haas, 1975; Haynes, 2002; Jones, 2008; Institute for the Advancement of Philosophy for Children [IAPC], 2002; Lipman, 1998; Russel, 2002).

6. Conclusion

P4C is relevant in fostering active and critical thinking as well as enhancing critical thinking (Green & Condy, 2016; Siddiqui, Gorard, & See, 2019). Therefore, P4C is feasible and practical to be implemented in the Moral

Education classroom because philosophy can provide a program that promotes thinking. However, there are several limitations to the study. Firstly, this study is limited in the sample population. The students were from two classes in the afternoon session and of Indian and Chinese ethnic groups, and the schools involved were chosen randomly by the researcher from schools in Gombak. Therefore, the findings from this study cannot be generalised to other different environments, subject, and duration.

Secondly, this study is constrained in terms of time. Time allocation of an hour per session for P4C in the Moral Education classroom was implemented. As a result, there was not enough time to complete the intervention of P4C in Moral Education during sessions one to three of the research. Therefore, the researcher had to cut short activities or exercises and moved to another slot due to the time constraint. The researcher also used the same stimulus for introduction and discussion for the next session. Instead of a text, images, videos and dilemmas were used as stimulus.

Based on the limitations of the research, the following recommendations are suggested. First, this research used a quasi-experimental design. Therefore, further study is recommended to apply action research. Action research is a process in which participants examine their own educational practise systematically and carefully, using the techniques of research. The use of action research is so that the teacher understands their students' existing problems and intent to solve either classroom problems or teaching problems.

Second, further research can be done and use a case study design to explore the teacher's views and the experiences in applying P4C. A case study is an in-depth investigation of a single person, group, event or community. By applying the case study, the researcher can provide rich qualitative information of P4C and explore the situational challenges and seek for the settlement of the issues. This is more meaningful by revealing the participants' real-life situations.

7. Acknowledgments

The authors express their appreciation to the Ministry of Education Malaysia (MOE), State Education Department, District Education Officers, the school principal, teachers and students for their involvement in this research. This research was supported by the Faculty of Education, National University of Malaysia code GG-2019-056.

References

- Allen, T. (1988). Doing philosophy with children. *Thinking: The Journal of Philosophy for Children*, 7(3), 23–28. doi: 10.5840/thinking19887322
- Ann, C. M. (1993). *Philosophy for children program and the development of critical thinking of gifted elementary students* (Doctoral dissertation, University of Kentucky, USA). Retrieved from <https://elibrary.ru/item.asp?id=5751571>
- Appoo, P. (2009). *The development and evaluation of moral reasoning module for the teaching and learning of moral education* (Doctoral dissertation). Universiti Sains Malaysia, Pulau Pinang.

- Balakrishnan, V. (2002). *Penggunaan dilemma real-life dalam Pengajaran Pendidikan Moral [Using real-life dilemma in teaching moral education]* (Master's thesis, Universiti Malaya, Kuala Lumpur). Retrieved from <http://pendidikanmoral.um.edu.my/uploads/rujukan/real%20life%20dilema.pdf>
- Balakrishnan, V. (2009). *Teaching moral education in secondary school students using real-life dilemmas* (Doctoral dissertation, Victoria University of Wellington, New Zealand). Retrieved from <http://researcharchive.vuw.ac.nz/xmlui/bitstream/handle/10063/1012/thesis.pdf?sequence=2>
- Barone, T. N. (2004). Moral dimension of teacher-student interactions in Malaysian secondary schools. *Journal of Moral Education*, 33(2), 179–196. doi:10.1080/0305724042000215221
- Berg, K., & Latin, R. (1994). *Essentials of modern research methods*. Englewood Cliffs, NJ: Prentice-Hall Inc
- Bono, E. D. (1995). *De Bono Thinking Course*. Great Britain: Clays Ltd, Stlves Plc.
- Burnes, B. (1981). *Philosophy for children. A report on achievement*. IAPC, Montclair State University. Retrieved from <http://cehs.montclair.edu/academic/iapc/research.shtml#cog>.
- Campbell, D. T., & Stanley, J. C. (1963). Experimental and quasi-experimental design for research on teaching. In Gage, N. (Ed.), *Handbook of research on teaching*, 171-246. Chicago: Rand McNally & Company.
- Che Abd Rahman, A. S. (2007). *Pengetahuan kandungan dan pedagogi guru pendidikan moral tingkatan empat di sebuah sekolah [Pedagogical content knowledge of moral education teacher among form four students in a school]* (Doctoral dissertation, Universiti Malaya, Kuala Lumpur, Malaysia). Retrieved from <http://repository.um.edu.my/642/1/Tesis%20Dr.%20Abd.%20Shatar%20bin%20Che%20Abd.%20Rahman.pdf>
- Creswell, J. W. (2005). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research* (2nd ed.). USA: Pearson.
- Creswell, J. W. (2009). *Research design, qualitative, quantitative and mixed methods approaches* (3rd ed.). Carlifonia: Sage.
- Christensen, L. B. (2004). *Experimental methodology* (9th ed.). Boston: Pearson.
- Dewey, J. (1938). *Experienced and education*. New York: Collier.
- Feuerstein, R., Falik, L. H., & Feuerstein, R. (2015). *Changing minds and brain- The legacy of Reuven Feuerstein: Higher thinking and cognition through mediated learning*. New York: Teacher College Press.
- Fisher, R. (2013). *Teaching thinking: Philosophical enquiry in the classroom* (4th ed.). London: Bloomsbury.
- Glickman, C. (1991). Pretending not to know what we know. *Educational leadership*, 48(8), 4–10. Retrieved from http://www.ascd.org/ASCD/pdf/journals/ed_lead/el_199105_glickman.pdf
- Green, L., & Condy, J. (2016). Philosophical enquiry as a pedagogical tool to implement the CAPS curriculum: Final year pre-service teacher perceptions. *South African Journal of Education*, 36(1). Retrieved from <http://dx.doi.org/10.15700/saje.v36n1a1140>
- Haas, H. J. (1975). Evaluation study: Philosophy for children. In M. Lipman, A. M. Sharp & F. Oscanyan (Eds.), *Philosophy in the classroom* (pp. 217–224). Philadelphia, PA: Temple University Press.
- Haberman, M. (1999). Increasing the number of high-quality African American teachers in urban schools. *Journal of Instructional Psychology*, 26(4), 208–212.
- Hashim, R. (2002). *P4C in Sri al-Huda Primary School* (Research report). Gombak: International Islamic University Malaysia.

- Hashim, R. (2013). *Pedagogi hikmah dan pengajaran kemahiran berfikir di sekolah [Wisdom pedagogy and critical thinking in school]*. Gombak: Centre for Philosophical Inquiry in Education, International Islamic University Malaysia.
- Hashim, R. (2017). P4C in the context of Muslim Education. In Naji, S., & Hashim, R. (Eds.), *History, theory and practice of philosophy for children: International perspectives* (pp. 170–179). London: Routledge. Retrieved from: <https://doi.org/10.4324/9781315208732-20>
- Hashim, R., Hussien, S., & Imran, A. M. (2014). Hikmah (wisdom) pedagogy and students' thinking and reasoning abilities. *Intellectual Discourse*, 22(2), 119–138. Retrieved from <http://journals.iium.edu.my/intdiscourse/index.php/islam/article/view/608>
- Hashim, R., Hussien, S., Khalid, M., & Wan Yusof, W. M. (2017). *Laporan kajian rintis program pedagogi hikmah di sekolah terpilih daerah Gombak: Bagi kemahiran berfikir kritis, kreatif, kolaboratif, komunikasi dan kepedulian [A pilot study of wisdom pedagogical program in selected schools in Gombak district: For critical thinking, creative, collaborative, communication and caring skills]*. (Research report), Gombak: Centre for Teaching Thinking, International Islamic University of Malaysia.
- Haynes, J. (2002). *Children as philosophers: Learning through enquiry and dialogue in the primary classroom*. London: Routledge.
- Hoon, C. L. (2010). An appraisal on the implementation of Moral Education for schools in Malaysia. *Proceedings of the 4th International conference on teacher education* (pp. 3–12). Bandung, Indonesia.
- Institute for the Advancement of Philosophy for Children. (2002). *IAPC research: Experimentation and qualitative information*. Retrieved from www.montclair.edu/pages/iapc/experimentalinfo.html
- Ivie, S. D. (1998). Ausubel's learning theory: An approach to teaching higher order thinking skills. *High School Journal*, 82(1), 35–42. Retrieved from <https://pdfs.semanticscholar.org/2ea6/7dc6c6a38c6425d6495f81d7e6902c8f79f2.pdf>
- Jones, H. (2008). Thoughts on teaching thinking: Perceptions of practitioners with a shared culture of thinking skills education. *Curriculum Journal*, 19(4), 309–324. doi:10.1080/09585170802509898
- Lam, C. M. (2012). Continuing Lipman's and Sharp's pioneering work on philosophy for children: Using *Harry* to foster critical thinking in Hong Kong students. *Educational Research and Evaluation: An International Journey on Theory and Practice*, 18(2), 187–203. Retrieved from <https://www.tandfonline.com/doi/abs/10.1080/13803611.2012.658669>
- Lim, T. K. (1994). Philosophy for children: Training of teachers for the Singapore programme. *Analytic Teaching*, 14(2), 53–56. Retrieved from <http://journal.viterbo.edu/index.php/at/article/view/605/392>
- Lipman, M. (1976). *Harry Stottlemeier's Discovery*. United States of America: Institute for the Advancement of Philosophy for Children, University Press of America Inc.
- Lipman, M. (1978). *Suki*. United States of America: Institute for the Advancement of Philosophy for Children, University Press of America Inc.
- Lipman, M. (1979). *Mark*. United States of America: Institute for the Advancement of Philosophy for Children, University Press of America Inc.
- Lipman, M. (1981). *Pixie*. United States of America: Institute for the Advancement of Philosophy for Children, University Press of America Inc.
- Lipman, M. (1982). *Kio and Gaus*. United States of America: Institute for the Advancement of Philosophy for Children, University Press of America Inc.
- Lipman, M. (1983). *Lisa*. United States of America: Institute for the Advancement of Philosophy for Children, University Press of America Inc.

- Lipman, M. (1987). *Elfie*. United States of America: Institute for the Advancement of Philosophy for Children, University Press of America Inc.
- Lipman, M. (1988). *Philosophy goes to school*. Philadelphia: Temple University Press.
- Lipman, M. (1993). *Thinking children and education*. Dubuque, Iowa: Kendall/Hunt Publishing Company.
- Lipman, M. (1996). Philosophical Discussion plan and exercises. *Analytic teaching*, 16(2), 64-77. Retrieved from <http://journal.viterbo.edu/index.php/at/article/view/635/422>
- Lipman, M. (1998). Teaching students to think reasonably; some findings of the philosophy for children program, *The Clearing House: A Journal of Educational strategies, Issues and Ideas*, 71(5), 277-280. Retrieved from <https://www.tandfonline.com/doi/abs/10.1080/00098659809602723?journalCode=vtch20>
- Lipman, M. (2017). The Institute for the Advancement of Philosophy for Children (IAPC) program. In Naji, S., & Hashim, R., (Eds.), *History, theory and practice of philosophy for children: international perspectives* (pp. 3-11). London: Routledge.
- Lipman, M., Sharp, A. M., & Oscanyan, F. S. (1980). *Philosophy in the classroom* (2nd ed.). Philadelphia: Temple University Press.
- Marashi, S. M. (2008). Teaching philosophy to children: A new experience in Iran. *Analytic Teaching*, 27(1), 12-15. Retrieved from <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.528.3963&rep=rep1&type=pdf>
- McLeod, T. L. (2010). *Philosophy for Children as a pedagogy for developing Oral English Language Skills with English a second language students* (Doctoral dissertation, New Mexico State University, Las Cruces, New Mexico). Retrieved from <https://search.proquest.com/openview/a2339fbc4d9241e326f0f01e0128a8ca/1?pq-origsite=gscholar&cbl=18750&diss=y>
- Minick, N. (2005). The development of Vygotskys' Thought: an introduction to thinking and speech. In Daniels, H (Eds.), *An introduction to Vygotsky*. (pp. 33-58). New York: Routledge.
- Ministry of Education. (2012). *Malaysia Education Blueprint 2013-2025: Preliminary report*. Putrajaya: Ministry of Education. Retrieved from <https://www.moe.gov.my/dasar/1207-malaysia-education-blueprint-2013-2025/file>
- Nachiappan, S., Sinnasamy, B., & Suffian, S. (2017). Masalah dalam pengajaran dan pembelajaran pendidikan moral di sekolah menengah dan penyelesaian masalah melalui kaedah hermeneutic [Problems in teaching and learning moral education in high school and problem solving through hermeneutic methods]. *Journal of Research, Policy & Practice of Teachers & Teacher Education (JRPPTTE)*, 7(2), 58-68. Retrieved from <http://ojs.upsi.edu.my/index.php/JRPPTTE/article/view/240>
- Othman, M. (2005). *Critical Thinking and reading skills: A comparative study of the reader response and philosophy for children program approaches* (Doctoral dissertation). Kuala Lumpur: International Islamic University Malaysia.
- Othman, M., & Hashim, R. (2006). Critical thinking & reading skills: A comparative study of the reader response & philosophy for children approaches. *Thinking: Journal of Philosophy for Children*, 8(2), 26-34. Retrieved from https://www.pdcnet.org/thinking/content/thinking_2006_0018_0002_0026_0034
- Pallant, J. (2007). *SPSS Survival Manual: A step by step guide to data analysis using SPSS for Windows, Version 15* (3rd ed.). Australia: Allen & Unwin.
- Piers, E. V. (1969). Manual for the Piers-Harris Children's Self-concept Scale. Nashville, TN: Counselor Recordings & Tescs. Retrieved from

- https://www.academia.edu/35084429/Piers-Harris_2_Piers-Harris_Childrens_Self-Concept_Scale_SECOND_EDITION
- Russell, J. (2002). Moral consciousness in a community of inquiry. *Journal of Moral Education*, 31(2), 141–153. doi:10.1080/03057240220143250
- Sasseville, M. (1994). Self-esteem, logical skills and philosophy for children. *Thinking: The Journal of Philosophy for Children*, 11(2), 30–32. doi:10.5840/thinking199411230
- Sharp, A. M. (2017). Philosophical novel. In Naji, S., & Hashim, R., (Eds.), *History, theory and practice of philosophy for children: international perspectives* (pp. 18-29). London: Routledge.
- Shipman, V. (1983). *New Jersey Reasoning Skills Test*. Upper Montclair, NJ: IAPC Test Division, Montclair State College.
- Siddiqui, N., Gorard, S., & See, B.H. (2019). Can programmes like philosophy for children help schools to look beyond academic attainment? *Educational Review*, 71(2), 146–165. Retrieved from <https://doi.org/10.1080/00131911.2017.1400948>
- Solomon, D., Watson, M. S., & Battistich, V. A. (2001). Teaching and schooling effects on moral/prosocial development. In V. Richardson (Ed.), *Handbook for research on teaching* (pp. 566–603). Washington, DC: American Educational Research Association.
- Yen, T. S., & Halili, S. H. (2015). Effective teaching of higher order thinking (HOT) in education. *The Online Journal of Distance Education and e-Learning*, 3(2), 41–47. Retrieved from <https://tojed.net/journals/tojed/articles/v03i02/v03i02-04.pdf>
- Zohar, A., & Schwartz, N. (2005). Assessing teachers' pedagogical knowledge in the context of teaching higher-order thinking. *International Journal of Science Education*, 27(13), 1595–1620. doi:10.1080/09500690500186592