ASCERTAINING THE CRITICAL THINKING AND FORMAL REASONING SKILLS OF STUDENTS'

Lim Tock Keng

Paper presented at the Growing Mind held in Geneva, Switzerland,
from 14-18 Sep 1996

ASCERTAINING THE CRITICAL THINKING AND FORMAL REASONING SKILLS OF STUDENTS[1]

Tock Keng Lim
National Institute of Education
Nanyang Technological University

ABSTRACT

With the critical thinking movement gaining momentum in all levels of education in the US and other countries, many thinking programmes have been developed. A thinking programme that emphasizes on process, teaching students *how* to think, rather than *what* to think, is the Philosophy for Children (P4C) programme, currently carried out in Singapore. A child, according to Matthew Lipman (1991), the founder of the P4C programme, could reason deductively and logically using concrete objects. In his specially written stories for children, Lipman translated the abstract formulations to reasoning in a concrete way that children could understand. To determine whether primary and secondary pupils in Singapore can reason and do philosophy, a study was set up in 1992 to ascertain their reasoning skills. Two instruments were used: the New Jersey Test of Reasoning (NJTR) specifically developed in the early 1980s to evaluate the P4C programme (Shipman, 1983) and the Test of Formal Reasoning (ATFR) written by Arlin (1982, 1984) to measure the stage of intellectual and cognitive level of the student: concrete, high concrete, transitional, low formal and high formal. This paper reports the findings of the study concerning the relationships between critical thinking as measured by the NJTR and concrete and formal reasoning as measured by the ATFR.

Introduction

Critical thinking is widely regarded as a generalized skill or ability (or a set of such skills and abilities) which can be utilized or applied across a variety of situations and circumstances. Lipman (1988) emphasized that if schools are to succeed in teaching critical thinking, educators must have a clear idea of what it is. To him, critical thinking is skilful, responsible thinking that facilitates good judgement because it relies upon criteria. Critical thinking is also self correcting

Paper presented at The Growing Mind, Geneva, Switzerland, September 14 - 18, 1996.

and is sensitive to context. To show that critical thinking instruction should emphasize on process teaching students *how* to think rather than *what* to think, Lipman developed the Philosophy for Children (P4C) programme, where he aimed to improve students' thinking by introducing them to philosophical issues embedded in specially written novels (for Grades K-12). Characters in the novels model the discovery of both formal and informal rules of thought. Students discuss philosophical issues through the passages from these novels (Lipman, Sharp & Oscanyon, 1980). The P4C programme is currently implemented successfully in a few Singapore schools (Lim & Koh, 1992; Lim 1994a, 1994b, 1996).

The P4C programme, one of the thinking programmes in the United States, is part of the critical thinking movement gaining momentum in all levels of education in many countries. Unlike some thinking programmes, it does not reduce thinking to being taught and mastered as a battery of atomic technical skills. In his novels for children used in the P4C programme, Lipman embedded abstract formulations, translates from what had been studied at university level philosophy, as concrete reasoning in stories of children understood by children; he feels that children could reason deductively and logically using concrete objects.

Many psychologists also consider that reasoning, like problem solving, is not necessarily a complex skill that only older children could do. For instance, Barell (1991) pointed out that Piaget's research on object permanence shows that at eighteen months babies begin to inquire about the object that, once in front of them, is placed under a rug. Paul, a leading exponent on critical thinking, argued that Piaget demonstrated that the thinking of young children presupposed philosophical foundations. To Paul (1990), most children have at least the impulse to philosophize and for a time seemed driven by a strong desire to know the basic *what* and *why* of things. However parents and teachers rarely cultivate this tendency. As children are usually given didactic answers in ways that discourage rather than stimulate thinking, they lose the impulse to question.

To determine whether primary and secondary pupils in Singapore can reason and do philosophy, a study was set up in 1992 to ascertain their reasoning skills. The study focused on the relationships between reasoning and inquiry as measured by the New Jersey Test of Reasoning (NJTR) specifically developed in the early 1980s to evaluate the Philosophy for Children (P4C) programme (Shipman, 1983), and concrete and formal operational reasoning as measured by the Test of Formal Reasoning (ATFR) written by Arlin (1982, 1984). The ATFR would measure the stage of intellectual and cognitive level of the student as being at one of five levels: concrete, high concrete, transitional, low formal and high formal. The study was conducted in 1992 on 160 Primary (Grade) 5 and 6 pupils from one school and 887 Secondary 1 and 2 (Grades 7 and 8) students from three schools.

Instruments

The NJTR consisted of 50 multiple choice items (with three options), each in the form of a short dialogue in simple language. It covered 22 skill areas of inductive and deductive reasoning and provided general information on critical thinking ability. Details of the reliability indices for the test were given in a final report on the P4C programme for an earlier 55-item version of the test; the Cronbach's alpha for samples from 5th to 7th grade classes ranged from .84 to .94 (Sutton, 1992).

To facilitate analysis in this study, 18 of the 22 skills (covered in 43 items) have been classified under 6 reasoning skills: recognising relationships (RR), inductive thinking (IT), evaluation (E), analysis (A), interpretation (I) and deductive thinking (DT). Recognizing relationships (RR1 to RR6) measures the recognition of symmetrical, transitive and causal relationships. The inductive thinking items (IT1 to IT6) assess the ability to come up with suitable and valid inferences based on given information and analogical reasoning. The 8 evaluation items (E1 to E8) involve the skill to identify good reasons and to assess the credibility of a statement from a source.

Analysis (A1 to A5) considers the detection of underlying and unstated assumptions together with the ability to weigh evidence to avoid unwarranted conclusions. Interpretation (I1 to I9) deals with both restating statements in alternative forms as well as expressing statements in logical form. Deductive thinking (DT1 to DT9) handles syllogistic reasoning, both in its categorical and conditional forms, as well as contradicting statements.

As reasoning skills of students might be somewhat constrained by the Piagetian stage of development that they were in, the ATFR was also used. The ATFR items were applications of Piaget's principles and not a direct translation of the Piagetian tasks. There were eight subtests measuring applications of Piaget's principles: volume, probability, correlation, combinations, proportions, momentum, mechanical equilibrium and frames of reference. The manual showed that the test-retest reliability of the subscales studied on a sample of 736 Grade 9 students were moderate ranging from 0.40 to 0.70 (Arlin, 1984). A multitrait multimethod validity study on 394 military recruits was also carried out by Arlin (1982).

The ATFR was designed, essentially, as a group test to measure the stage of intellectual and cognitive level of the student - concrete, high concrete, transitional, low formal and high formal. Pupils in the concrete and high concrete level would be in Piaget's concrete operational stage while pupils in the transitional stage would exhibit some instances of formal reasoning (Inhelder & Piaget, 1958). Pupils in the low and formal stage represented performance of Piaget's formal reasoning tasks.

Analyses and Results

The data on the two instruments, the NJTR and ATFR, collected from the four schools were analyzed. Table 1 set out the Piagetian levels (as identified by the ATFR) of the students by schools. A is a good primary school, B is an average secondary school while C and D are good secondary schools. As expected, the primary pupils were mainly in the concrete (36.9%) and high concrete (56.3%) stages. The lower secondary students of the average school, B, were mainly in the high concrete (65.2%) stage while half the lower secondary students in the good schools were mainly in the low formal stage. The ATFR appeared to classify the students of the four schools in their respective Piagetian cognitive stages.

Insert Table 1 about here

Using SAS Version 6 (SAS Institute Inc., 1990), a one-way analysis of variance (ANOVA) was used to determine whether the mean scores of the NJTR (the entire test of 50 items) differed significantly between the 5 cognitive levels of the ATFR. The Student-Neuman-Keuls (SNK) test was also performed to identify where the differences lie. The mean NJTR scores of each of the 5 cognitive stages, as presented in Table 2, differed significantly ($F = 217.46$, $p < .001$). The majority of pupils in the primary school, who were mainly on the concrete and high concrete level, had a mean of 23.88 and 34.38 respectively. As the total score of the NJTR was 50 it would appear that even primary pupils at the concrete stage were able to answer about half the reasoning questions of the NJTR. This indicated that primary level pupils appeared to be able to reason.

Insert Table 2 about here

To determine the types of reasoning skills that primary students could handle at the concrete and formal Piagetian stages, the reduced NJTR test with 6 main reasoning skills was used. The concrete stage consists of 384 students in levels 1 and 2 while the formal stage includes 454 students in levels 4 and 5. Rasch analysis (one model of Item Response Theory) was carried out on the 43 NJTR items in each of the two stages, using the program, Quest (Adams & Khoo, 1996). The Quest program uses the joint or UCON maximum likelihood procedure to estimate both the item and person parameters with a correction factor for bias.

The model fit for the reduced NJTR test in the concrete and formal stages was analyzed using the item and case infit and outfit statistics reported in Quest; these were the weighted and unweighted residual-based statistics described by Wright and Masters (1982). The reduced NJTR test in the concrete stage had a infit mean square of 0.99 (with SD of 0.15) and in the formal stage had an infit mean square of 0.98 (with SD of 0.06); both data sets therefore fit the model. The reliability of estimate is also good, 0.97 to 0.98 for both tests. This estimate is the Wright and Masters' item separation reliability for the proportion of the observed estimate variance that is considered "true".

The variable map for the concrete stage (see Figure 1) shows the distribution of the cases (students) and the items along the reasoning skill variable. The logit scale for the calibration of items and cases are plotted vertically on the map such that the distribution of the person ability level can be compared with the distribution of the item difficulty level. The variable maps for the concrete (Figure 1) and formal (Figure 2) stages were then analyzed to ascertain the level of the reasoning skills in the students.

Insert Figure 1 about here

Insert Figure 2 about here

In the variable map of the concrete stage in Figure 1, the distribution of the sample (N = 384) appears to skew slightly to the right, ie. towards higher ability. This is shown in the summary statistics of reasoning skills and cases, presented in Table 3; the 384 cases ranged from an ability level of -2.91 logits to 1.91 logits, with a mean of 0.62 logits and a standard deviation of 0.83 logits while the 43 items ranged from a difficulty level of -1.54 logits (RR3) to 1.94 logits (I6). the mean of the items, in accordance with the Rasch model, was centred at 0.00 and the standard deviation was 0.93.

Insert Table 3 about here

The higher reasoning ability of the students in the formal level was indicated in the variable map of Figure 4. The cases all appear to cluster in the top left quadrant, extending from -1.70 logits to 4.64 logits with a mean of 2.56 and a standard deviation of 0.82. The items, on the other hand, are spread along the left side of the variable map, with a minimum value of -2.71 (RR1, RR3 and IT6) and a maximum value of 2.71 (I6), the mean being centred at 0.00 with a standard deviation of 1.52.

The mapping at Table 4 compared clearly the distribution of the items and students. At the concrete level, close to 100% of the students matched the logit level of all the items, even though the items clustered at a lower logit level than the cases. The contrast was, however, much greater at the formal level, where only about 68% of the students ability level matched the difficulty level of about 54% of the items; the remaining items were far too simple for the students.

Insert Table 4 about here

A study of the difficulty level of the 6 reasoning skills is displayed in Table 3. Both the concrete and formal level students found the skill RR of recognising relationships to be the simplest. The concrete level showed a mean of -0.69 and S.D. of 0.47 while the formal level has a mean of -1.19 and S.D. of 1.36) levels. The most difficult skill appeared to be DT, deductive thinking; the concrete level has a mean of 0.57 and S.D. of 0.86 whereas the formal level indicated a mean of 0.92 and S.D. of 1.41.

Conclusion

The NJTR has provided useful information of the reasoning ability of Singapore students, particularly at the Piagetian concrete and formal levels. The reasoning level of primary and secondary students in Singapore showed that it was appropriate to introduce the P4C programme. The four schools in the Singapore P4C programme are doing well.

In terms of research on the NJTR, there is a need to further investigate the relationships between biographical factors, such as age and gender, with reasoning. Home-language speaking environment is another important factor in Singapore, as students are taking a test in a non-native language. Even though schools in Singapore use English as the medium of instruction, not all students speak English at home. Students tend to use their mother tongue, Chinese, Malay or Tamil at home. The relationship between such biographical factors and reasoning skills will help educators to fine tune thinking programmes for Singapore students.

REFERENCES

Adams, R. J. & Khoo, S. T. (1996) Quest: the interactive test analysis system, Version 2.1, Hawthorn, Victoria: ACER.

Arlin, P. K. (1982) A multitrait-multimethod validity study of a test of formal reasoning, Educational and Psychological Measurement, 42, 1077 - 1088.

Arlin, P. K. (1984) Arlin Test of Formal Reasoning, East Aurora, NY: Slosson.

Barell, J. (1991) Teaching for thoughtfulness: Classroom strategies to enhance intellectual development, New York: Longman.

Lim T. K. (1994a) Philosophy for children: Training of teachers for the Singapore programme, Analytic Teaching, 14, 2, 53 - 56.

Lim T. K. (1994b) The Philosophy for Children Project in Singapore, Thinking, 11, 2, 33 - 37.

Lim T. K. (1996) Fostering critical thinking in the classroom, Korean Journal of Thinking and Problem Solving, 6, 1, 63 - 74.

Lim, T. K. & Koh S. K. (1992) Philosophy for Children: Experiences in two schools, Proceedings of the 6th Annual Conference of the Educational Research Association, Singapore

Lipman, M. (1988) Critical Thinking -- What can it be, Educational Leadership, 46, 1, 38 - 43.

Lipman, M., Sharp, A. M., & Oscanyon, F. S. (1980) Philosophy in the Classroom, Philadelphia: Temple University Press

Inhelder, B. & Piaget, J. (1958) The growth of logical thinking from childhood to adolescence, New York: Basic Books.

Paul, R. (1990) Critical Thinking: What every person needs to survive in a rapidly changing world, Rohnert Park: C.A.: Centre for Critical Thinking and Moral Critique.

SAS Institute Inc., (1990) SAS User's Guide: Statistics (Version 6), Cary, NC: SAS Institute.

Shipman, V. (1983) The New Jersey Test of Reasoning Skills, Totowa, NJ: Totowa Board of Education.

Sutton, R. E. (1992) Review of the New Jersey Test of Reasoning, in J. J. Kramer & J. C. Conoley, Eleventh Mental Measurement Yearbook, Lincoln, NE: Burros Institute of Mental Measurement, page 908.

Wright, B. D. & Masters, G. (1982) Rating Scale Analysis: Rasch measurement, Chicago: MESA Press

Table 1 Frequency of Students by Schools and Piagetian Level of Reasoning

Level*	School A	School B	School C	School D
Concrete	59 (36.9%)	6 (4.4%)	1 (0.2%)	1 (0.3%)
High Concrete	90 (56.3%)	90 (65.2%)	75 (17.1%)	62 (20.0%)
Transitional	9 (5.6%)	23 (16.7%)	110 (25.1%)	67 (21.6%)
Low Formal	2 (1.2%)	18 (13.0%)	218 (49.6%)	156 (50.3%)
High Formal	0 (0.0%)	1 (0.7%)	35 (8.0%)	24 (7.8%)
Total	160 (100%)	138 (100%)	439 (100%)	310 (100%)

* A is a primary school while B, C, and D are secondary schools.

Table 2 Means of New Jersey Test of Reasoning scores by Formal Level as measured by the Arlin Test of Formal Reasoning

Level	N	Mean	SD	F value
Concrete	67	23.88	8.69	217.46***
High Concrete	317	34.38	8.32	
Transitional	209	40.94	4.52	
Low Formal	394	42.99	3.87	
High Formal	60	45.08	2.94	

*** $p < .001$

Figure 1 Person-Ability Item-Difficulty Map for NJTR items (N=43) for Students (N = 384) at the Concrete Level

```
------------------------------------------------------------------------------------------------
Ability Distribution                        Item Difficulty Distribution
------------------------------------------------------------------------------------------------
 5.0                         |
                             |
                             |
                             |
                             |
                            X|
 4.0                         |
                             |
                             |
                            X|
                             |
 3.0                      XXX|
                           XX|
                        XXXXX|
                    XXXXXXXXX|
 2.0              XXXXXXXXXXX|                                    I6
                             |                                                          DT1

                      XXXXXXX|                                               A5
                 XXXXXXXXXXXX|
                  XXXXXXXXXXX|               IT3          E6
                  XXXXXXXXXXX|                                                          DT2
 1.0                  XXXXXXX|                                               A2         DT5,DT6
                  XXXXXXXXXXX|
       XXXXXXXXXXXXXXXXXXXXXX|                            E7                            DT3
                   XXXXXXXXXX|
                   XXXXXXXXXX|                            E2,E5
                      XXXXXXX|                                        I3     A1         DT4
               XXXXXXXXXXXXXX|               IT1                                        DT9
 0.0                    XXXXX|               IT2,IT5                  I1
                        XXXXX|                            E8          I5
                    XXXXXXXXX|   RR5
                        XXXXX|   RR4,RR6                                                DT8
                          XXX|                                        I4,I9
                           XX|   RR2                                  I7
                       XXXXXX|   RR1         IT6          E1,E3
-1.0                        X|               IT4                      I2     A3         DT7
                           XX|
                           XX|                                               A4
                           XX|                            E4          I8
                            X|   RR3
                            X|
                             |
-2.0                         |
                             |
                             |
                             |
                             |
                            X|
-3.0                         |
------------------------------------------------------------------------------------------------
```

Each X represents 2 students

Figure 2 Person-Ability Item-Difficulty Map for NJTR items (N=43) for Students (N = 454) at the Formal Level

```
------------------------------------------------------------------------------------------------------------------
Ability Distribution                                      Item Difficulty Distribution
------------------------------------------------------------------------------------------------------------------
5.0                              |
                                 |
                         XXXXX   |
                                 |
                                 |
                                 |
                                 |
4.0                              |
                     XXXXXXXXX   |
                                 |
                                 |
                  XXXXXXXXXXXX   |
                                 |
         XXXXXXXXXXXXXXXXXXXXX   |
3.0                              |
          XXXXXXXXXXXXXXXXXXXX   |
                                 |                                        I6                       DT1
               XXXXXXXXXXXXXXX   |
                                 |
            XXXXXXXXXXXXXXXXXX   |                                                    A2
                XXXXXXXXXXXXXX   |                             E6
2.0                XXXXXXXXXXX   |                                                    A5           DT2,DT6
                                 |
                        XXXXXX   |                             E2                                  DT3,DT5
                        XXXXXX   |
                         XXXXX   |
                            XX   |
1.0                          X   |
                                 |                             E5         I5
                                 |                                        I1                       DT9
                                 |                             E7
                                 |                IT1                     I2
                                 |   RR5
                                 |   RR6          IT3                                 A1
0.0                              |                                        I3,I7                    DT4
                                 |
                                 |                IT2          E8
                                 |                IT5                     I9
                                 |
                                 |   RR2
                                 |
-1.0                             |                             E3
                                 |                             E4                                  DT7
                                 |                                                                 DT8
                                 |
                                 |   RR4          IT4                     I4          A3
                                 |                                                    A4
                                 |
-2.0                             |
                                 |                             E1         I8
                                 |
                                 |
                                 |
                                 |   RR1,RR3      IT6
                                 |
-3.0                             |
------------------------------------------------------------------------------------------------------------------
  Each X represents     3 students
```

Table 3 Summary Statistics of the Reasoning Skills and Cases in Logits

		Concrete Level				Formal Level			
Reasoning Skill / Cases	N	Mean	S.D.	Min.	Max.	Mean	S.D.	Min.	Max.
Recognising Relationships (RR)	6	-0.69	0.47	-1.54	-0.29	-1.19	1.36	-2.71	0.30
Inductive Thinking (IT)	6	-0.01	0.82	-0.93	1.34	-0.68	1.23	-2.71	0.56
Evaluation (E)	8	-0.02	0.92	-1.40	1.30	0.10	1.45	-2.20	2.02
Interpretation (I)	9	-0.18	0.95	-1.35	1.94	0.11	1.43	-2.20	2.71
Analysis (A)	5	0.17	1.26	-1.25	1.61	0.21	1.86	-1.71	2.22
Deductive Thinking (DT)	9	0.57	0.86	-0.96	1.83	0.92	1.41	-1.29	2.63
Cases (Students in C/F)	384/454	0.62	0.83	-2.91	1.91	2.56	0.82	-1.70	4.64

Table 4 Item Difficulty and Student Ability Mapping at the Concrete and Formal Levels

Item Difficulty/ Person Ability (logits)	Concrete Level				Formal Level			
	Items	%	Cases	%	Items	%	Cases	%
$-3.01 \le L \le -2.0$	-	-	2	0.3%	5	11.6%	-	-
$-2.01 \le L \le -1.0$	4	9.3%	14	3.6%	8	18.6%	1	0.2%
$-1.01 \le L \le 0$	19	44.2%	67	17.4%	7	16.3%	-	-
$0.01 \le L \le 1.0$	11	25.6%	175	45.6%	13	30.2%	5	1.1%
$1.01 \le L \le 2.0$	9	20.9%	126	33.0%	6	14.0%	97	21.3%
$2.01 \le L \le 3.0$	-	-	-	-	4	9.3%	206	45.3%
$3.01 \le L \le 4.0$	-	-	-	-	-	-	129	28.5%
$4.01 \le L \le 5.0$	-	-	-	-	-	-	16	3.6%
Total	43	100%	384	100%	43	100%	454	100%