CHAPTER FOUR

RESULTS AND DISCUSSION

This chapter presents both the statistical results and the discussion according to the research questions of this study. First, descriptive statistical treatments were adopted to display the distribution of six question types, and their percentage. Second, the results of the participants' performances in the pretest were analyzed and discussed to identify their problems and to reveal whether there is any significant difference between the control and the experimental groups before the formal reading instruction. Finally, after the formal reading instruction, the participants took both of the BCT in 2006 and the results of their performances were gathered and analyzed. The statistical differences between the control and the experimental groups were presented and discussed to display the effectiveness of the reading instruction.

Research Question (1)

What are the question types and their frequency distribution in the BCT?

Results of the Question Types and Their Percentage

In the first section, descriptive statistics was performed to investigate the number and percentage of the six individual reading question types not only in the reading pretest but also in the reading tests of the BCT in 2006.

The Reading Pretest

The materials of the reading pretest were adopted from the BCT from 2002 to 2004, including 21 reading passages and 55 reading questions. According to the classification from Johnson (2004), Mo (1987) and the BCT committee (2005), there

are totally six question types in the reading tests: identifying the main idea (MI), finding the detailed information (DI), determining the meaning out of the context (MC), finding the targets of the reference (TR), drawing implications and inferences (II), and drawing correct conclusions (CC). The frequency distribution, percentage and ranking of these six question types in the pretest are presented in Table 1.

Table 1

The Frequency distribution, Percentage and Ranking of Question Types in the Pretest

Question Type			Frequency	Percentage	Rank
	DI		27	49.1	1
	TR		8	14.5	2
Literal skills	MC		5	9.1	4
		DI + TR + MC	40	72.7	
	MI		5	9.1	4
	II		6	10.9	3
Critical skills	CC		4	7.3	6
		MI +II + CC	15	27.3	
Total			55	100	

Notes: 1. The literal reading skills include DI, TR and MC while the critical reading skills include MI, II and CC (Johnson, 2004).

2. DI+ TR+ MC refers to the sum of the three question types related to the literal skills, and MI + II + CC refers to the sum of the other three question types for the critical skills.

Table 1 shows that for the reading tests in the BCT from 2002 to 2004, the question type DI is the most frequent one with almost fifty percent distribution, followed by the question type TR (14.5%), II (10.9%), MI and MC (9.1%), and the

least frequent question type is CC (7.3%).

The frequency distribution of these six question types in the pretest showed an unbalanced allocation among the 55 reading questions. Nearly half of these questions were assigned to the question type DI. However, less than 10 percent of the reading questions belong to the question types MI, MC and CC respectively. It is reasonable that a reading passage may contain more questions on the DI type because a certain number of supporting details could be included in the reading passage to make the main idea clear and to complete the flow of the writer's thoughts and opinions. However, the proportion of this question type which put too much emphasis on the basic reading skills of matching the questions with detailed information could be cut down to allow the increment of other question types such as MI and CC that may involve more on deeper reading comprehension and reading strategies.

Johnson (2004) further divided the reading skills into two major categories, the literal and the critical reading skills. The question types DI, TR and MC are the basic and literal skills "because they depend on simple awareness and understanding of the ideas that are stated (p. 64)." On the other hand, the critical reading skills include the question types MI, II and CC, which may involve the skills to "recognize the author's purpose, the overall organization, or relationships within and between sentences (p. 172)."

Farr et al. (1990) mentioned the types of questions that follow a reading selection will determine if the reading selection focuses on only the surface meaning of the text or on other- perhaps deeper – comprehension. Hughes (1989) also suggested the inclusion of both the micro-skills and macro-skills for the reading comprehension tests. Johnson (2004) even focused more on the critical reading skills that may help develop readers' reading ability such as making inferences, identifying the main idea and drawing conclusions. In the pretest, however, the situation of

unbalanced distribution between the literal and critical reading skills is obvious. The percentage of the question types belonging to the literal reading skills is over 70 percent and those for the critical reading skills are not equally emphasized.

Huang (2004) indicated that the BCT have created both positive and negative backwash effects on EFL teaching. Junior high school English teachers provide more authentic materials similar to the BCT for extensive reading but focus more on basic reading skills and strategies than on the critical ones which may enhance students' understanding overall and essential parts of the reading passages such as the main idea, the conclusion, and the writer's purpose and intention. Because the reading tests of the BCT have emphasized too much the question type finding the detailed information, teachers may be influenced to provide instruction to meet this tendency; thus, they often neglect the need to help improve learners' critical reading skills and strategies.

The BCT committee have brought up junior high school English teachers' attention to take up the reading instruction (You, 2004); however, they may try hard to think of striking a balance among different reading question types and reading skills to enhance students' overall reading ability. Teachers are supposed to enhance learners' critical reading skills as well as the basic ones. Therefore, this research focused not only on providing a more complete reading instruction in junior high school English class, but also on suggesting the authorities develop a more balanced design of various reading question types and important reading skills which may be beneficial to the learners' overall reading ability.

The BCT in 2006

The participants took both reading tests of the BCT in 2006, the first one in May and the other in July. The first English reading test consisted of 6 passages, 16 questions and the other included 7 passages, and 17 questions. In comparison to those

in the reading pretest, there are only five reading question types: identifying the main idea (MI), finding the detailed information (DI), determining the meaning out of the context (MC), finding the targets of the reference (TR), and drawing implications and inferences (II). The question type drawing correct conclusions (CC) was not recognized in either of the BCT in 2006. The frequency distribution, percentage and ranking of these question types are presented in Table 2.

Table 2

The Frequency distribution, Percentage and Ranking of Question Types
in the BCT 0f 2006

Question Type		Frequency			Percentage			Rank		
		F	S	F+S	F	S	F+S	F	S	F+S
	DI	9	10	19	56.3	58.7	57.5	1	1	1
Literal	MC	3	1	4	18.7	5.9	12.1	2	4	3
skills	TR	3	3	6	18.7	17.7	18.2	2	2	2
	DI + MC + TR	15	14	29	93.7	82.3	87.8			
	MI	1	0	1	6.3	0	3.1	4	6	5
Critical	II	0	3	3	0	17.7	9.1	6	2	4
skills	CC	0	0	0	0	0	0	6	6	6
	MI + II + CC	1	3	4	6.3	17.7	12.2			
Total		16	17	33	100	100	100			

Notes: 1. The literal reading skills include DI, TR and MC while the critical reading
skills include MI, II and CC (Johnson, 2004).
2. DI+ TR+ MC refers to the sum of the three question types related to the
literal skills, and MI + II + CC refers to the sum of the other three question
types for the critical skills.
3. F refers to the first BCT of 2006 in May while S refers to the second BCT
of 2006 in July. F+S refers to the combination of both tests in 2006.

Table 2 incorporates both reading tests of the BCT in 2006 and the question type DI still holds the strongest position in the reading tests, and even exceeds its percentage in the pretest. In the first BCT of 2006, there are 16 reading questions and 9 of them belong to this question type (56.3%). There are 3 questions categorized as the question type MC (18.7%) and another 3 questions assigned to the question type TR (18.7%). Only one question is classified as the question type MI (6.3%). But no reading question is found to fit into the question types II and CC.

In the second BCT in July, there are 17 reading questions and there are 10 reading questions identified as the question type DI (58.7%). There are also 3 questions belonging to the question type TR (17.7%) and another 3 reading questions are classified as the question type II (17.7%). There is only one question assigned to the question type MC (5.9%). As for the question types MI and CC, no reading question is found.

In general, the DI type is still the most dominating question type in both BCT of 2006 (57.5%). Questions types TR and MC exist in both of the BCT and they rank second and third (18.2%, 12.1%). Question types II and MC appear only in one of the BCT and they rank fourth and fifth (9.1%, 3.1%). And no reading question is recognized in either BCT as the question type CC.

To be more specific, in both of the BCT, the question types related to the literal reading skills have an overwhelming percentage over the critical ones, 93.7 % to 6.3% and 82.3 % to 17.7%. The focus is mostly on the question types that are related to the literal reading skills (87.8%), while the question types of the critical reading skills are not equally emphasized (12.2%). Moreover, the question types of the literal reading skills are regularly observed in both of the tests, while those of the critical reading skills appear in only one of the tests or even none of them. These may result in the unbalanced distribution of different question types and the reading skills required.

In the pretest, the distribution of different question types seems unbalanced; however, this unbalanced distribution becomes even worse in both of the BCT in 2006. This unbalanced situation of over-emphasis on the literal skills may bring about the negative backwash effect. The question type DI becomes the most dominating one, and teachers and students may focus mostly on this basic and literal reading skills. The critical reading skills may be ignored and learners may lack the ability of deeper comprehension (Farr et al., 1990). Johnson (2004) also mentioned "the literal comprehension skills enable you to determine what it is that the author has said on the page. But *really* understanding written material takes more than just being able to identify ideas that have been stated" (p. 172). For the reading questions of the BCT in 2006, too many reading questions are about finding the detailed information and other literal skills that may emphasize the surface meanings of the texts. Besides, in comparison with the results in the pretest, the question types MI, and II are cut down and this may lead to another negative backwash effect that teachers and students may not pay much attention to these critical reading skills and strategies required in processing the reading passages. Moreover, the question type CC does not appear in either of the BCT in 2006, and this may result in teachers' and students' ignorance of this reading skill in their reading instruction or learning, and students may not be asked to undergo deeper comprehension or processing of the reading passages, and their reading ability in these aspects may not be emphasized and developed.

Therefore, researchers considered it important to include not only the literal reading skills but also the critical ones into the reading comprehension tests. For example, Dubin et al. (1986) suggested that reading skills be addressed and items be designed in a way that these reading skills are evaluated. Yu (1995) pointed out that an ideal reading test should provide items that enable readers to demonstrate their reading ability. Nuttal (1996) also encouraged the test constructor to define the

reading skills and to write questions which involve such skills.

However, the reading tests of the BCT in 2006 did not seem to be designed according to these criteria. Students' literal reading skills such as finding the detailed information were tested in a large proportion. They may rely strongly on such reading skills and look for only the surface meanings. Other critical reading skills may be rarely or even not required when students answered the reading questions. For example, in the researcher's school, there are ten English teachers. After a short interview with all of them, six of them never taught certain reading skills to their students. Another three have taught reading skills in class but not the critical ones. They considered it a waste of time because the BCT did not put much emphasis on such question types and reading skills. Consequently, these important reading skills may not be evaluated and students' related reading ability may not be demonstrated. All of these problems may lead to serious negative backwash effects in the reading tests of the BCT. Therefore, it is recommended that the committee reconsider the distribution and proportion of different reading question types to enhance teachers and students' awareness of these critical reading skills and strategies as well as the literal ones.

In sum, the reading tests of the BCT in 2006 may not be a very good model for the reading instruction because the question types were not well–designed and well-distributed. Therefore, the reading tests of the BCT may leave room for criticisms and discussions, and the BCT committee may take the negative backwash effect mentioned above into consideration and try to rectify such negative effect the BCT has brought into the English class. There is no doubt that the BCT may lead the reading instruction, but what's more important, it has to lead the instruction in the positive and active way.

Research Question (2)

How do the participants perform in each question type and what are their problems?

Results of the Participants' Performances in the Pretest

The participants took the reading pretest during the class in February of 2006, and they were told to take this pretest as a formal exam which may demonstrate their reading comprehension ability and problems. One and a half hours of the testing time were given to each of the participants to complete all the reading questions. There are totally 60 students who completed the pretest, 30 in the experimental group and another 30 in the control group. As for those students who could not complete the pretest, their performances were not taken into account. There are 13 students who could not complete the reading pretest, six in the control group and seven in the experimental group. According to Hughes (2000), when participants' overall ability is tested, guessing may have a considerable effect on test scores and the reliability of the test may be threatened. Bernhardt (1991) also distinguished already literate from non-literate L2 learners who were not treated as a homogeneous group. The students who could not complete the pretest may be considered as non-literate L2 learners. Thus, it is suggested that their performances be excluded from the statistic data.

Hsu (2004) also mentioned that the major difficulties junior high school students confronted when reading English texts were a lack of vocabulary, grammatical knowledge and effective reading strategies. The researcher then interviewed each of these students to figure out what their problems were. Most of them reported that they did not know what the questions were about and a lot of words in the reading passages were unfamiliar to them. And they failed to complete the reading pretest. Even if they had finished the test, the results would have been

unreliable. Therefore, for those who failed to complete the pretest, the results of their

performances were excluded to strengthen the reliability of the test.

      For the results of the participants' performances in the control and the

experimental groups in the pretest, *t*-test and one-way ANOVA were processed to

analyze their performances. The results of the pretest may serve as an important

indicator to see if there is any significant difference between these two groups before

the formal reading instruction. Table 3 is the results of the participants' overall

performances which include the mean scores, the percentage, standard deviation and

standard error mean of the two classes in the pretest.

Table 3

The Participants' Overall Performances in the Pretest

Class	N	Mean	Percentage	Standard Deviation	Standard Error Mean
A	30	39.37	71.58	11.99	2.19
B	30	39.20	71.27	13.16	2.40
Total	60				

Notes. 1. Class A refers to the control group while Class B refers to the experimental
group.
    2. The number of the participants in both classes is 30, and those who failed to
complete the pretest were not included.
    3. The maximal mean is 55.

      According to Table 3, the mean scores, which refer to the average performance

of all the reading questions in the pretest, are similar in Class A and Class B. The

mean score in Class A is 39.37 while it is 39.20 in Class B. That is, out of the 55

reading questions, the participants in Class A answered correctly for 39.37 questions

(71.58%) and in Class B 39.20 questions (71.27%). The standard deviation in Class A

is 11.99 and 13.16 in Class B. Besides, the standard error mean in Class A is 2.19 and

2.40 in Class B. The participants' overall performances in the experimental and the

control groups were very close to each other, and there seemed to be little difference

between these data. However, statistic treatments have to be processed and to make

sure if there is any significant difference between these two groups before the formal

reading instruction. The result of the *t*-test processing of the participants' overall

performances in the pretest is shown in Table 4.

Table 4

T-test Processing of the Participants' Overall Performances in the Pretest

t	df	Sig.(2-tailed)	Mean Difference	Standard Error Difference
.051	58	.959	.17	3.251.

Note. The mean difference is significant at the .05 level.

The result of the t-test processing of the mean scores shows that there is no

significant difference in the participants' overall performances in the pretest

between these two classes (t=.051, p=.959), which indicates that the participants from

the two different classes performed alike and made no significant difference on their

performances. This may ensure the condition that before the formal reading

instruction, the participants' overall reading performances are similar in both groups.

Though there is no significant difference observed for their overall performances, a

further investigation of the differences for individual question types is taken into

consideration. Table 5 presents the means scores of individual question types between

the control and the experimental groups.

Table 5

Mean Scores of Individual Question Types in the Pretest

Question Type		Class A	Percentage	Rank	Class B	Percentage	Rank
	DI	19.53	72.3	2	19.37	71.7	2
Literal	MC	3.37	67.4	3	3.27	65.4	4
skills	TR	6.00	75.0	1	5.80	72.5	1
	DI + MC +TR	28.9	72.2		28.44	71.1	
	MI	3.34	66.8	4	3.29	65.8	3
Critical	II	3.10	51.7	6	3.17	52.8	6
skills	CC	2.43	60.6	5	2.56	64.0	5
	MI + II + CC	8.87	59.1		9.02	60.1	

Notes: 1. The literal reading skills include DI, TR and MC while the critical reading skills include MI, II and CC (Johnson, 2004).
2. DI+ TR+ MC refers to the sum of the three question types related to the literal skills, and MI + II + CC refers to the sum of the other three question types for the critical skills.
3. The total number for DI is 27, MC is 5, TR is 8, MI is 5, II is 6, and CC is 4.

The results of the mean scores for individual reading question types are also similar in Class A and Class B. For example, for the question type DI, the participants in Class A answered 19.53 questions correctly out of 27 questions (72.3%), while those in Class B answered 19.37 questions (71.7%). When it comes to the question type MI, the mean score in Class A is 3.34 (66.8%) and in Class B 3.29 (65.8%). For the question type MC, it is 67.4% in Class A, 65.4% in Class B. The mean scores and percentage are similar to one another in the two classes for the rest of the question types. Moreover, the ranking of these six different question types in the pretest also

presents similar tendency. The question type TR seems the easiest one for the participants in both groups. The question types DI ranked second. In Class A, the question type MC ranks superior to the question type MI (67.4 % to 66.8%). But the ranking is reversed in Class B (65.4% to 65.8%). However, the results of these two question types are very close between Class A and Class B. For the question types CC and II, the participants in both classes did not perform very well and their ranks are five and six.

To be more specific, the participants performed better in the question types concerning the literal reading skills (72.2% & 71.1%) than in those concerning with the critical ones (59.1 & 60.1). This is probably because such question types related to the critical reading skills may require not only the surface meanings but also the deeper processing and comprehension of these meanings. According to Johnson (2004), the three literal comprehension skills are usually the easiest skills to learn because they depend on simple awareness and understanding of the ideas that are stated. They merely require attention to the ideas the author has stated. On the contrary, the critical skills need more than just the surface meanings, and understanding the author's intentions is often the crucial element of successful comprehension in related question types. Therefore, when taking reading comprehension tests, students may have more problems with the question types related to the critical reading skills such as MI, II and CC.

Moreover, Hughes (2000) also proposed the micro-skills which are similar to the literal reading skills and the macro-skills are like the critical ones. The micro-skills are recognized as the exercise of straight-forward grammatical and lexical abilities, but the macro-skills are related either to needs or to course objectives. The learners often acquire the micro-skills earlier and easier. So, before the formal reading instruction, the participants in both groups performed better on the question

types related to the literal reading skills than those of the critical ones.

To sum up, the participants' performances on the question types of the literal or micro-skills are better than those of the critical or macro-skills, and they may have more trouble with the question types concerning the critical reading skills. The results show that the control and the experimental groups performed alike in individual question types, and the problems they have are also similar. However, it is important to figure out whether there is any significant difference among these individual question types. Table 6 presents the results of the participants' performances in individual question types with one-way ANOVA.

Table 6

The Participants' Performances of Individual Question Types in the Pretest

	Source	*SS*	*df*	*M S*	*F*	*Sig.*
	DI	2156.85	59	37.18	.01	.92
Literal	MC	129.98	59	2.19	.07	.79
skills	TR	225.40	59	3.88	.16	.70
	Sum L	4645.3	59	3.27	.41	.84
	MI	132.40	59	2.28	.00	.87
Critical	II	110.93	59	1.91	.04	.85
skills	CC	98.18	59	1.69	.25	.62
	Sum C	753.65	59	.42	.03	.86

Notes: 1. The literal reading skills include DI, TR and MC while the critical reading
        skills include MI, II and CC (Johnson, 2004).
    2. Sum L refers to the sum of the three question types related to the literal
        skills, and Sum C refers to the sum of the other three question types for the
        critical skills.

In Table 6, one-way ANOVA was processed to make sure if there is any individual difference among these six different question types in Class A and Class B. For example, for the question type DI, there seems no significant difference between the two groups (F= .01, P= .92). The same result is calculated for other five question types: MC (F= .07, P= .79), TR (F= .16, P= .70), MI (F= .00, P= .89), and CC (F= .28, P= .62). Moreover, the question types related to the literal (F= .41, P= .84) and the critical (F =.03, P= .86) reading skills did not present significant difference, either. The results in Table 6 show that there is no significant difference between the control and the experimental groups for individual question types and the question types related to the literal and the critical reading skills.

To sum up, the results of the reading pretest display that there exists no significant difference not only for the participants' overall performances but also for their performances on the six individual question types. Therefore, the participants' performances in the experimental group seem similar to those of the participants in the control group before the formal reading instruction.

Research Question (3)

How do the reading skills and strategies work in class for junior high school participants?

Results of the Participants' Performances in the BCT of 2006

The participants in the control and the experimental groups took both of the BCT held in May and July of 2006. There are about 45 testing items in the English test and about one third of these items can be classified as the reading comprehension questions. In the following section, the results of the participants' performances on the English reading comprehension questions in the BCT of 2006 were analyzed to see

the differences of their overall performances between the control and the experimental

groups, and to check if there is any significant difference among the six individual

question types, and between the literal and the critical reading skills after the formal

reading instruction. Both descriptive and inferential statistic treatments were

conducted in the present research to process and analyze the participants'

performances. In Table 7, the participants' overall performances in the first and the

second BCT of 2006, including the mean scores, the percentage, standard deviation

and standard error mean, are presented and compared.

Table 7

The Participants' Overall Performances in the BCT of 2006

Class	BCT 2006	N	Mean	Percentage	Standard Deviation	Standard Error Mean
A	F	30	14.43	90.19	1.89	.36
	S	30	15.67	92.18	2.38	.44
B	F	30	12.50	78.10	3.54	.67
	S	30	14.03	82.53	4.08	.76

Notes. 1. F means the BCT held in May and S means the BCT in July. Class A refers
          to the experimental group while class B refers to the control group.
       2. The number of the participants in both classes is 30, and those who failed to
          complete the pretest were not included.
       3. The maximal mean of the reading questions in the first BCT is 16, and the
          maximal mean of the reading questions in the second BCT is 17.

For the first BCT of 2006, the participants in the experimental group performed

better than those in the control group, and the mean score is 14.43 (90.19%) to 12.5

(78.1%). The results of the second BCT of 2006 are similar to those of the first one;

however, the difference seems to reduce, 15.67 (92.18%) to 14.03 (82.53%). In general, the participants in Class A performed better than those in Class B for both of the BCT in 2006. Those who are from the experimental group answered the reading questions correctly more than 90 percent of the questions, but those from the control group answered correctly about 80 percent of the questions.

In contrast to the results of the participants' overall performances in the pretest, which displayed few differences, the results of the participants' overall performances in the BCT of 2006 after the formal reading instruction seemed to display greater differences between these two groups, and the experimental group performed better than the control one. To be more specific, the data collected from both of the BCT in 2006 must be processed by means of the statistic treatments of *t*-test to identify whether there is any significant difference among of the participants' overall performances. Table 8 displays the *t*-test treatment of their overall performances in the first and the second BCT of 2006 between the experimental and the control groups.

Table 8

T-test Processing of the Participants' Overall Performances in the BCT of 2006

BCT-2006	t	df	Sig.	Mean Difference	Standard Error Difference
F	3.23*	44.27	.002	2.37	.73
S	1.89	46.72	.065	1.63	.86

Notes: 1. F means the first BCT in May and S means the second in July.
     2. * p < .05.

Table 8 presents the t-value, the mean difference, standard error difference and the significance of the participants' overall performances in the first and the second BCT of 2006. The mean difference is significant for the first BCT of 2006 (t= 3.23,

p= .002) but the difference for the second BCT of 2006 is not significant

(t= 1.89, P=.065). This means that the formal reading instruction for the experimental

group is significant and effective in the first BCT of 2006. After the formal reading

instruction, the participants who learned the reading skills and strategies in the

experimental group performed better than those who practiced the reading passages

one after another in the control group.

On the other hand, the result also shows that there is no significant difference in

the second BCT of 2006, which may result from the time factor; the first BCT was

held in May but the second one was in July. Mo (1987) suggested that the reading

skills and strategies be practiced and reinforced over and over again to help learners

acquire and internalize these skills and strategies. Brown (1994) also encouraged the

development of different reading strategies and suggested the learners to use them.

However, the participants in the researcher's school graduated in June and after their

graduation, they did not go back to school for any formal instruction. For the first

BCT of 2006 held in May, both the experimental and the control groups had a lot of

practice before the test, and the former focused on the reading skills and strategies

while the latter emphasized the content-based instruction in the passage-by-passage

process. The results in the first BCT of 2006 may be more reliable than those in the

second one because after the participants graduated, the researcher could not actually

guide their practice but ask them to have as much practice as possible on their own.

The effectiveness of the formal reading instruction may reduce and the results of the

participants' overall performances in the second BCT of 2006 may, therefore, become

not significant. Generally speaking, the results of the participants' overall

performances in the first BCT of 2006 are significant and the formal reading

instruction seems effective for the experimental group.

Moreover, there is a need to identify the effectiveness and significance of the

formal reading instruction on the six individual question types and Table 9 presents

the mean scores of individual question types in the first BCT of 2006.

Table 9

Mean Scores of Individual Question Types in the first BCT of 2006

Question Type		Class A	Percentage	Rank	Class B	Percentage	Rank
	DI	7.87	87.44	4	6.93	77.00	3
Literal	MC	2.83	94.33	2	2.43	81.00	2
skills	TR	2.90	96.67	1	2.57	85.67	1
	DI + MC +TR	13.60	90.67		11.93	79.53	
	MI	.90	90.0	3	.57	57.0	4
Critical	II						
skills	CC						
	MI + II +C C	.90	90.0		.57	57.0	
Total		14.43	90.19		12.50	78.1	

Notes: 1. The literal reading skills include DI, TR and MC while the critical reading
            skills include MI, II and CC (Johnson, 2004).
         2. DI+ TR+ MC refers to the sum of the three question types related to the
            literal skills, and MI + II + CC refers to the sum of the other three question
            types for the critical skills.
         3. The total number for DI is 9, MC is 3, TR is 3, and MI is 1 in the first BCT
            of 2006.

In Table 9, the means scores of individual question types show that the

participants in the experimental group performed better on the four question types

appearing in the first BCT of 2006. For example, the mean score of the question type

DI is 7.87 (87.44%) in Class A, but 6.93 (77.00%) in Class B, the question type MC is

2.83 (94.33) in Class A, but 2.43 (81.00%) in Class B, the question type TR is 2.90 (96.67%) in Class A, but 2.57 (85.67%) in Class B, and the question type MI is .90 (90.00%) in Class A, but .57 (57%) in Class B. In the first BCT of 2006, there is no reading question classified as the question type II or CC.

The participants' performances in individual question types displayed greater differences in the first BCT of 2006 than in the reading pretest between the experimental and the control groups. For instance, the mean difference for the question type DI in the pretest is 0.16 (0.60%) while it is 0.94 (10.44%) in the first BCT of 2006, the difference for the question type MC is 0.1 (2.00%) in the pretest and 0.4 (13.33%) in the first BCT of 2006, TR is 0.2 (2.50%) in the pretest and 0.33 (11.00%) in the first BCT of 2006, and MI is 0.05 (1.00%) in the pretest and 0.33 (33.00%) in the first BCT of 2006. The mean differences for the participants' performances of individual question types between the two groups in the first BCT of 2006 are over 10 percent and for the MI type, it is even over 30 percent. Therefore, the differences between the experimental and the control groups become greater for the individual question types in the first BCT of 2006. The participants from the experimental group performed better than those from the control group after the formal reading instruction.

In addition, the mean difference of the question types related to the literal reading skills also displays greater difference in the first BCT of 2006 (11.14%) than in the pretest (1.10%). As for the question types that can be classified as the critical reading skills, the mean difference is even greater in the first BCT of 2006 (33.00%) than in the pretest (1.00%). In the pretest, there are few differences of the mean between the literal and critical skills, but in the first BCT of 2006, the mean difference is more than 20 percent. The result shows that after the formal reading instruction, the participants in the experimental group may become more capable of applying the

reading skills to tackle their reading comprehension questions, especially the critical ones. However, there may be a problem to come to this conclusion because of the limitation of unbalanced distribution among different question types in the first BCT of 2006. As shown in Table 9, there is only one reading question categorized as the question type MI and no reading question can be identified as the question type II or CC. Thus, there is only one reading question which belongs to the critical reading skills and the result may not be representative and reliable enough due to the problems of the sample number.

On the other hand, the ranking of individual question types in the first BCT of 2006 is compared to that in the pretest. First of all, the question type TR still ranks the highest in both groups in the first BCT of 2006. Secondly, the question type MC ranks higher than the DI type in the first BCT of 2006, but this ranking is reversed in the pretest. Thirdly, in the pretest, the question types included in the literal skills rank higher than those of the critical skills in both the experimental and the control groups. However, in the first BCT of 2006, the percentage of the mean score and ranking of the question type classified as the critical skills display great differences between the experimental and control groups. For the experimental group, the percentage and ranking of the question type MI are similar to those of the literal skills, but the percentage and ranking of the same question type are much lower than those of the literal skills in the control group. This result again implies that after the formal reading instruction, the participants in the experimental group may be able to apply the critical reading skills as well as the literal ones; but for the participants in the control group, their performances on the question types of the critical reading skills are still much worse than those of the literal ones. This may suggest that the instruction of the critical reading skills is effective in the experimental group.

To sum up, after the formal reading instruction, the participants' performances

in the first BCT of 2006 are greatly different from those in the pretest. The

participants in the experimental group performed well in both the question types of

the literal and the critical skills and better than the performances out of the control

group. The performances of the critical skills improved in the experimental group but

not in the control group. Greater differences were observed in the first BCT of 2006;

however, there is a need to further verify the significance of the differences for

individual question types between the two groups in the first BCT of 2006 with the

statistic processing of one-way ANOVA.

Table 10

The Participants' Performances of Individual Question Types in the First BCT of 2006

	Source	*SS*	*df*	*M S*	*F*	*Sig.*
	DI	246.40	59	13.07	3.25	.077
Literal	MC	21.93	59	2.40	7.13*	.010
skills	TR	19.73	59	1.67	5.35*	.024
	Sum L	472.73	59	41.67	5.61*	.021
Critical	MI	12.60	59	1.07	5.36*	.024
skills	Sum C	12.60	59	1.07	5.36*	.024

Notes: 1. The literal reading skills include DI, TR and MC while the critical reading
skills include MI, II and CC (Johnson, 2004).
2. Sum L refers to the sum of the three question types related to the literal
skills, and Sum C refers to the sum of the other three for the critical skills.
3. * p < .05.

In Table 10, there are significant differences among the question types MC (F=

7.13, P= .10), TR (F= 5.35, P= .024) and MI (F= 5.36, P= .024). But the result from

the question type DI is not significant (F= 3.25, P= .077). Moreover, the results from

the sum of the literal reading skills and the critical ones are both significant (F= 5.61

P= .021; F= 5.36, P= .024). This may indicate that for the participants in the experimental group, the formal reading instruction of the reading skills and strategies may be effective not only on the individual question types but also on the literal and critical reading skills.

The results presented in Table 10 suggest that the formal reading instruction of the reading skills and strategies may be helpful among the three individual question types MC, TR and MI. The participants' performances in these three question types in the experimental group are significantly higher than those in the control group, and the reading instruction for the question types MC, TR and MI is effective.

Besides, the results also show that both the literal and the critical reading skills in the formal reading instruction may be effective for the participants in the experimental group. But in the reading pretest (see Table 5), the participants in both groups performed better on the question types of the literal reading skills than the critical ones. In the experimental groups, the participants' performances on the critical reading skills are similar to the literal ones, while for those in the control group, their performances on the critical skills are still much worse than the literal ones, just like the results in the pretest. Again, this may imply that the instruction of the question types related to the critical skills is effective in the first BCT of 2006. However, this conclusion may not be reliable because only one reading question was observed to be the question type of the critical skills.

On the contrary, the result from the question type DI shows no significant difference between the two groups in the first BCT of 2006. This means the instruction of the reading skills and strategies about the question type DI is not effective enough to display any significant difference between the experimental and the control groups. According to Johnson (2004), the literal skills are much easier for learners than the critical ones. Because the question type DI belongs to the literal

skills, the answers to this question type can be easily figured out if the participants have enough time to read the passage carefully. Most of the participants also responded that the time for the BCT English test is long enough for them to check their answers at least once. Besides, among the three question types of the literal skills, finding the detailed information is probably the one required the least skills. The participants only had to understand what the questions were about and matched the questions with the lines in the reading passage. But for the other two question types, they had to guess the meaning from the contextual clues or decide on the possible candidates of the reference. Therefore, the participants in the control group could not figure out the answers for the question types MC and TR as easily as the question type DI, and the participants' performances in the experimental group with the instruction of related reading skills and strategies may present more effective results in this two question types than in the question type DI. This may explain why the result from the question type DI is not significant but the other two question types also included in the literal skills can be significant.

In addition, for the question type MI, which belongs to the critical reading skills, the participants had to first identify the topic sentence and then paraphrase the sentence into its title or the main idea for the reading questions. Hughes (2000) also thought of the micro-skills (literal skills) as means of promoting the development of the macro-skills (critical skills). Therefore, the question types concerning the literal reading skills may go with some basic skills and strategies, while those related to the critical ones may need more advanced skills and strategies. As discussed above, because the answers to the question type which requires more advanced skills can not be easily figured out by the participants in the control group, the participants in the experimental group, with the instruction of the reading skills and strategies to identify the main idea, may perform better than those in the control group.

Moreover, the results that indicated the effectiveness of the instruction of the reading skills and strategies in the first BCT of 2006 could be referred back to the literature review on the theory of reading. In interactive model, reading involves both the applications of higher mental operation and the lower text processing. Both bottom-up and top-down approaches are important elements to complete the reading tasks. According to Criper and Davies (1988), reading skills are text-oriented as bottom-up model, while reading strategies are reader-oriented as the top-down model. Therefore, the instruction of the reading skills and strategies are helpful for the participants in the experimental group to complete their reading tasks. Also, Coady (1979) suggested a three-component model, which involves the interaction of the reader's "conceptual abilities, background knowledge, and process strategies (p. 7)." The reading skills and strategies are similar to the component, process strategies, and the effectiveness of the formal reading instruction may help verify the importance of this component.

To sum up, in the first BCT of 2006, the results from the participants' performances of individual question types MC, TR and MI are significantly different between the experimental and the control groups and the sums of the results in the literal and critical reading skills are also significant. Therefore, after the formal reading instruction, the participants in the experimental group did perform better on these three individual reading question types than the participants in the control group.

On the other hand, the results of the participants' performances in the second BCT of 2006 may present different outcomes because it was held in July, which was about one and a half months after the first BCT. Moreover, the participants graduated from the researcher's school in June. After that, they did not go back to school for any formal instruction. Though it is shown in Table 8 that there is no significant difference for the participants' overall performances in the second BCT of 2006, the

effectiveness and significance of individual question types in this test still need to be discussed and compared. Table 11 presents the mean scores of individual question types in the second BCT of 2006.

Table 11

Mean Scores of Individual Question Types in the Second BCT of 2006

Question Type		Class A	Percentage	Rank	Class B	Percentage	Rank
	DI	9.40	94.00	1	8.60	86.00	1
Literal	MC	.83	83.33	4	.67	66.67	4
skills	TR	2.7	90.00	3	2.5	83.33	2
	DI + MC +TR	12.93	92.36		11.77	84.07	
	MI						
Critical	II	2.73	91.00	2	2.27	75.67	3
skills	CC						
	MI + II +C C	2.73	91.00		2.27	75.67	

Notes: 1. The literal reading skills include DI, TR and MC while the critical reading skills include MI, II and CC (Johnson, 2004).
2. DI+ TR+ MC refers to the sum of the three question types related to the literal skills, and MI + II + CC refers to the sum of the other three question types for the critical skills.
3. The maximal mean for DI is 10, MC is 1, TR is 3, and II is 3.

In Table 11, just like the first BCT of 2006, the participants' performances on individual question types for the experimental group are better than those for the control group. For example, the participants in Class A answered 9.40 questions (94.00%) correctly among 10 reading questions classified as the question type DI, but the participants in Class B answered only 8.60 questions (86.00%) correctly out of the

same questions. For the question type MC, the mean score of the experimental group

is .83 (83.00%), but that of the control group is .67 (67.00%). The mean score of the

question type TR is 2.7 (90.00%) in Class A, but 2.5 (83.33%) in Class B. For the

question type II, the participants in the experimental group got the mean score of 2.73

(91.00%) and those in the control group got 2.27 (75.67%). Therefore, the participants

in the experimental group performed better on these four individual question types

than those who were in the control group. As for the question types MI and CC, no

reading question was observed in the second BCT of 2006.

In addition to the individual differences between the four reading question types

discussed above, the question types categorized as the literal and the critical reading

skills also display mean differences between the experimental and the control groups.

The mean score of the literal skills is 12.93 (92.36%) in class A, but it is 11.77

(84.07%) in Class B. For the critical reading skills, only the question type II is

identified in the second BCT of 2006, and due to the limitation of unbalanced

distribution of the reading question types, the results of the mean difference from the

critical reading skills may become less representative and reliable than those from the

literal ones. However, the results of the mean differences from both the literal and the

critical reading skills in the second BCT of 2006 present greater difference than those

in the pretest.

When the results in the second BCT of 2006 are compared with those in the

first BCT, it is found that the results of the participants' performances on individual

question types does not reveal greater mean differences between the experimental and

the control groups. For example, the mean differences of the question type DI are .94

(10.44%) in the first BCT of 2006, and .80 (8.00%) in the second BCT of 2006. For

the question type TR, the mean differences are .33 (11%) in the first BCT of 2006,

and .20 (6.67%) in the second BCT of 2006. But the mean difference of the question

type MC becomes even greater in the second BCT of 2006. It is .16 (16.66%), but .4 (13.33%) in the first BCT of 2006. Moreover, greater difference than the results in the pretest are found concerning the question type II, which belongs to the critical reading skills. Unlike those of the literal skills, the participants in the control group performed better on this question type in the pretest and the mean difference in .07 (1.1%), but in the second BCT of 2006, the participants in the experimental group performed much better on the same question type and the mean difference is .46 (15.33%).

Generally speaking, the participants in the experimental group still performed better than those in the control group in the second BCT of 2006, but the results also show that the mean differences reduced in contrast to the participants' performances in the first BCT of 2006. According to Mo (1987) and Brown (1994), the effectiveness of the reading skills and strategies may become less prominent because of time and lack of practice. As mentioned earlier, the participants graduated from the researcher's school about one and a half months before the second BCT of 2006 and some of them may feel more relaxed and even forget to review their reading skills and strategies. In such cases, the effectiveness of the formal reading instruction may become less significant and this may also threaten the reliability of the second BCT of 2006 in the present research.

In addition to the mean differences discussed above, there are also differences in the ranking of individual question types that the participants performed among the pretest, the first and the second BCT of 2006. The participants in both groups answered the reading question TR the best not only in the pretest but also in the first BCT of 2006, and they also performed better on the question types of the literal reading skills than those of the critical ones, the ranking of the question types which belong to the literal skills are usually higher than those related to the critical ones (see Table 5 and Table 9). However, in the second BCT of 2006, the ranking of the

question types in Table 11 presents a different order. The participants' performances

on the question type DI rank the top of four individual question types, and the

question type MC ranks the last instead of the question type II. In the experimental

group, the question type II even ranks second. The ranking of individual question

types in the second BCT of 2006 again may not be very reliable because there is only

one reading question identified as the question type MC. Besides, item difficulty of

the reading questions is not analyzed in the present research and this may affect

greatly on the ranking of different question types.

In sum, the results from the second BCT of 2006 indicated that the participants

in the experimental group performed better than those in the control group, but the

differences between both groups reduce when the results are compared with those in

the first BCT of 2006. The participants in the experimental group performed well not

only on the question types of the literal reading skills but also on those of the critical

ones, but the participants in the control group performed better on the literal skills

than on the critical ones. Moreover, the ranking of four individual question types in

the second BCT of 2006 is different from the ranking in the pretest and in the first

BCT of 2006. Although the differences between the experimental and the control

groups reduce and the participants' overall performances in the second BCT of 2006

are not significant between Class A and Class B (see Table 8), the statistic treatment of

one-way ANOVA can be processed to reveal the significance of individual question

types in the second BCT of 2006.

Table 12 displays the significance of four individual question types between the

experimental and the control groups in the second BCT of 2006. Among these

question types, only the result from the question type II is significant (F= 4.14,

p= .046). The other three question types which are included as the literal reading skills

are not significant (F= 2.84, p= .098; F=2.23, p= .141; F= .921, p= .341), and the sum

of the question types related to the literal skills is not significant, either (F= 2.89, p= .094). As mentioned earlier, the participants in both groups graduated in June but the second BCT was held in July, and the effectiveness of the formal reading instruction may become less significant.

Table 12

The Participants' Performances of Individual Question Types in the Second BCT of 2006

	Source	SS	df	M S	F	Sig.
	DI	206.00	59	9.60	2.84	.098
Literal	MC	11.25	59	.42	2.23	.141
skills	TR	38.40	59	.60	.921	.341
	Sum L	429.65	59	20.42	2.89	.094
	MI					
Critical	II	49.00	59	3.27	4.14*	.046
skills	CC					
	Sum C	49.00	59	3.27	4.14*	.046

Notes: 1. The literal reading skills include DI, TR and MC while the critical reading skills include MI, II and CC (Johnson, 2004).
2. Sum L refers to the sum of the three question types related to the literal skills, and Sum C refers to the sum of the other three question types for the critical skills.
3. * p < .05.

Moreover, though the results from the question types of the literal skills are not significant, the question type II which belongs to the critical reading skills is still significant. This may bring about another finding that the instruction of the critical reading skills may be more effective than those of the literal ones. Since the

effectiveness of the formal reading instruction on the question types of the literal skills has decreased in the second BCT of 2006, the result of the literal skills still presents significant difference between the experimental and the control groups. According to Johnson (2004), the question types of the literal skills are easier for learners than the critical ones, and the participants can figure out the answers of the reading questions classified as the literal skills more easily. When the effectiveness of the formal reading instruction decreases, the significance of the literal reading skills may be reduced as well because the participants in the control group may come out with correct answers without these reading skills if they are given enough time to finish their reading questions.

On the other hand, the instruction of the critical skills remains significant because even if the effectiveness of the formal reading instruction decreases, the participants in the experimental group can still figure out the answers with their reading skills but those in the control group can not come out with correct answers as easily as the literal ones. However, this finding may need to be verified with further research because of insufficient reading questions that can be identified as the question types of the critical reading skills.

All in all, the results of the present research has shown that first, there are totally six difference question types for the reading questions in the BCT, and they are DI, MC, TR, MI, II and CC. But the question type CC can not be identified in either the first or the second BCT of 2006. Moreover, the first three question types can also be categorized as the literal reading skills while the other three are the critical skills. The results also show that the frequency and percentage of different question types which belong to the literal and the critical skills display an unbalanced distribution not only in the pretest but also in the first and the second BCT of 2006. There seems to be too much emphasis on the question types of the

literal reading skills. Second, the results of the participants' performances in the pretest suggest that they performed better on the question types of the literal reading skills than the critical ones. Third, the results show that the formal reading instruction is effective in the participants' overall performances and their performances on individual question types in the first BCT of 2006. Finally, the results of the participants' performances in the second BCT of 2006 reveal that the effectiveness of the formal reading instruction with reading skills and strategies may decrease because of the practicing time and efforts. Besides, the results in the second BCT of 2006 also suggest that the instruction of the critical reading skills may be more effective than that of the literal ones but this finding may not be very reliable because of the limited number of the sample questions. Therefore, this needs to be testified with further research that may include more reading questions of the literal reading skills.