Newcastle University

Infusing Thinking Skills into an L2 Classroom: a case study of an innovation in a Taiwanese university

MEI-HUI CHEN

Doctor of Philosophy

School of Education, Communication and Language Sciences

Newcastle University

November 2010

Declaration

I hereby declare that this thesis is my own work and effort and that it has not been submitted previously for a degree at this or any other university. Any material or sources of information used in this study have been fully acknowledged.

Signature: ..... Mei Hui Chen

Date: ....Nov. 2010.....

Acknowledgements

阿　彌　陀　佛

Amitabha

I am very grateful to my supervisors, Dr Kate Wall and Dr Mei Lin, for their guidance, advice, patience and encouragement. In particular, I am indebted to Dr Wall who trained me well to think critically on a doctoral level. I would also like to show my appreciation to Professor David Leat for his help with designing the thinking tasks. Lots of thanks also go to Alice and Barbara who assisted in the process of analysing the data and our lecturers in School of ECLS who gave me advice for the study.

I feel an immense gratitude to my family, Mum, Dad, my brothers and sister-in-laws, especially, to my lovely daughter's understanding and support. Without their full backing, I would surely not have been able to complete this thesis. In these five years, they have always been there for me, given me strength and comforted me when I felt down. In particular, my Dad would generously support me with English pounds while I was in England, and my Mum would kindly contribute her New Taiwan dollars when I was back home in Taiwan.

In addition, I would also like to thank my good friends in Taiwan and my colleagues in the study room 2.41 for their continuing friendship. I will never forget those numerous freezing cold nights staying in the study room and fighting with my thesis, and my colleagues coming to me to encourage me to keep going on.

Finally, a big thank you also goes to Amy and the participants for taking part in this study. Despite this study being especially cognitively demanding, they cooperated and took the challenge willingly.

Regardless of all the romantic notions of pursuing a PhD, most of this journey, I would say, was pretty tough, challenging and sometimes upsetting. It was, however, a very fruitful journey and I enjoyed it.

Abstract

Empirical research on higher-order questions has shown positive impact on student achievement in L1 and L2 classrooms in western countries and former British colonies. In association, a world-wide increase in teaching thinking has led to a debate about its applicability to L2 settings, especially to Asian learners. In Taiwan speaking has been identified as a problematic area in L2 learning and alongside equipping students with thinking skills has been highlighted as a goal of Higher Education (HE). Therefore, this study undertook a case study of an innovation where Higher-Order Thinking Skills (HOTS) approach was used to try and enhance students' L2 speaking proficiency and thinking skills. This approach had two steps. In a teacher-led setting the innovator modelled the tasks to the students which focused on how to answer higher-order questions, interact and comment on others' opinions. Students then practiced the thinking tasks in groups as part of their L2 learning. The impact on speaking and thinking and the effect on students' attitudes were examined to determine the practicability of this approach in a Taiwanese university L2 classroom. Two classes of non-English major freshmen participated in this study: one class received the innovation, while the other class did not. A mixed-method approach was applied and data collected in three phases: pre-, post- and delayed post-test. In contrast to Atkinson's (1997) claim that using a critical thinking pedagogy to teach non-native speakers of English in L2 classrooms is inappropriate, the findings show strong evidence to support the idea that a HOTS approach enhances learners' speaking and thinking performance with the majority of students holding positive attitudes. This indicates infusing thinking skills into the L2 classroom is practicable and students can be trained as active thinkers. A most significant finding was the occurrence of *high-cognitive interactive talk,* which created numerous opportunities for speaking and thinking. This tackled the L2 speaking problems observed and met the goals of HE, i.e. it equipped university students with thinking skills and encouraged active learning.

Table of contents

DECLARATION I

ACKNOWLEDGEMENTS II

ABSTRACT III

TABLE OF CONTENTS IV

FIGURE VIII

TABLE IX

TERMINOLOGY X

CHAPTER 1 - EXPLORING THE RESEARCH CONTEXT 1

1.1 Introduction 1

1.2 Outlining the issues 1

1.3 Exploratory study 2

1.3.1 Analysis and findings 2

1.3.2 Discussion 5

1.4 Context for the study 7

1.5 The HOTS approach 11

1.5.1 Design of thinking tasks 12

1.5.2 Lesson plan 15

1.5.3 Implementation 16

1.6 Purpose of the study and research questions 18

1.7 Value of this study 19

CHAPTER 2 - LITERATURE REVIEW 20

2.1 Introduction 20

2.2 Theories of learning 20

2.2.1 Social constructivism 20

2.2.2 Theories of second language acquisition 24

2.2.3 Summary: exploring optimal conditions for L2 learning 27

2.3 Thinking skills 28

2.3.1 What are thinking skills? 28

2.3.2 Thinking skills programmes in L1 classrooms 29
2.3.3 Thinking skills in L2 classrooms 30
2.3.4 The teaching and learning of thinking 32
2.4 Bloom's taxonomy 34
2.4.1 Lower-order thinking 35
2.4.2 Higher-order thinking 36
2.4.3 Difficulty and complexity 38
2.5 Higher-order questions 39
2.5.1 Formulating higher-order questions 40
2.5.2 The debate about the effectiveness of higher-order questions 42
2.5.3 Higher-order questions in L2 classrooms 43
2.6 Strategies relevant to questioning 47
2.6.1 Wait-time 47
2.6.2 Probing 49
2.7 Theoretical conceptual framework 50
2.7.1 The rationale 51
2.7.2 An explanation of the framework 52
CHAPTER 3 - METHODOLOGY 55
3.1 Introduction 55
3.2 Aims and research questions 55
3.3 A critique of methods used in the field 56
3.4 Methodological basis 58
3.5 Research design 63
3.5.1 A mixed-method approach 63
3.5.2 A case study of an innovation 66
3.5.3 An overview of data collection procedures 67
3.6 Participants and sampling 68
3.7 Justification of data collection tools 72
3.7.1 Video recording 72
3.7.2 Questionnaires 73
3.7.3 Interviews 76
3.8 Analysis 79
3.8.1 Video recording 80

3.8.2 Interviews 89
3.8.3 Questionnaires 89
3.9 Reliability and validity 91
3.9.1 Reliability 91
3.9.2 Validity 93
3.10 Ethical issues 96
3.11 Summary 98
CHAPTER 4 - RESULTS 99
4.1 Introduction 99
4.2 The effect on speaking 100
4.2.1 Perceptions of opportunities for speaking 100
4.2.2 Perceptions of the effect on speaking 102
4.2.3 The effect on speaking 104
4.2.4 Summary 112
4.3 The effect on thinking 113
4.3.1 Perceptions of the opportunities for thinking 113
4.3.2 Thinking in action 115
4.3.3 The effect on thinking 120
4.3.4 Perceptions on the ability and using thinking skills 126
4.3.5 Summary 132
4.4 Attitudes and perceptions 133
4.4.1 Attitudes towards and perceptions of the HOTS approach 136
4.4.2 Perceptions of the effect on students' performance 143
4.4.3 Evaluation on innovation tools 149
4.4.4 Summary 153
CHAPTER 5 - DISCUSSION 155
5.1 Introduction 155
5.2 Contexts for socio-cognitive conflict 157
5.2.1 Design and implementation of thinking tasks 158
5.2.2 Thinking tasks as contexts for socio-cognitive conflict 162
5.2.3 Exploring patterns of interaction 166
5.3 High cognitive interactive talk 169
5.4 Performance outcomes 172

5.4.1 Cognitive cultivation 172
5.4.2 L2 speaking development 180
5.4.3 Learning behaviour 192
5.4.4 Attitudes and perceptions 194
5.4.5 Reflecting on L2 speaking problems in Taiwan 198
5.5 Summary 199

CHAPTER 6 - CONCLUSION 203

6.1 Introduction 203
6.2 Limitations of the study 203
6.3 Implications for HE 205
6.4 Recommendations for teachers 206
6.5 Suggestions for future research 207
6.6 Significance of this study 210

BIBLIOGRAPHY 213

APPENDICES 238

Figure

Figure 1. Quadrants of linguistic and cognitive demands in tasks.....12
Figure 2. Odd One Out.....14
Figure 3. Make Up A Story.....14
Figure 4. Guess What I Say.....15
Figure 5. Bloom's taxonomy of six thinking levels.....35
Figure 6. Theoretical conceptual framework.....53
Figure 7. Methodological conceptual framework.....59
Figure 8. A comparative case study design.....66
Figure 9. Median of study group members' average score.....71
Figure 10. A framework of data analysis and interpretation.....80
Figure 11. Mean turn length.....105
Figure 12. The total number of words.....106
Figure 13. The total number of turns.....107
Figure 14. Individual mean turn length in the innovation class.....108
Figure 15. Grammatical complexity.....108
Figure 16. The total number of clauses.....109
Figure 17. The total number of AS units.....110
Figure 18. Individual grammatical complexity in the innovation class.....110
Figure 19. Overall speaking proficiency.....111
Figure 20. Individual speaking proficiency in the innovation class.....112
Figure 21. Median of speaking proficiency in the innovation class.....112
Figure 22. Lower-order and higher-order thinking skills used by the innovation class.....121
Figure 23. Six thinking skills used by the innovation class.....121
Figure 24. Lower-order and higher-order thinking skills used by the comparison class.....122
Figure 25. Six thinking skills used by the comparison class.....122
Figure 26. The function of HOTS used by the innovation class.....124
Figure 27. The function of HOTS used by the comparison class.....125
Figure 28. The attitudes towards the HOTS approach.....136
Figure 29. Perceptions of the effect of the HOTS approach on students' performance.....146
Figure 30. Descriptive analysis of attitudes towards the thinking tasks.....150
Figure 31. Quadrants of linguistic and cognitive demands of the thinking tasks.....150
Figure 32. The overall structure for discussion.....156
Figure 33. Patterns of peer interaction.....167
Figure 34. High cognitive interactive talk wheel.....170

Table

Table 1. An example of weekly lesson plan 16
Table 2. Data collection procedures.....68
Table 3. The categories of questionnaires 90
Table 4. The priority of data used for analysis 100
Table 5. Analysis of One-way ANCOVA on the comparison of the perceived ability to answer lower-order questions 127
Table 6. Analysis of One-way ANCOVA on the comparison of the perceived ability to answer higher-order questions.....127
Table 7. Paired-Samples T Test on the perceived ability to answer lower-order questions.....128
Table 8. Paired-Samples T Test on the perceived ability to answer higher-order questions.....128
Table 9. Mean scores of the perceptions of the ability to answer higher-order questions.....129
Table 10. Analysis of One-way ANCOVA on the comparison of the perceived importance of using lower-order thinking.....130
Table 11. Analysis of One-way ANCOVA on the comparison of the perceived importance of using higher-order thinking 130
Table 12. Mean score of the importance of using higher-order thinking.....130
Table 13. Paired-Samples T Test on the perceived importance to use lower-order thinking 131
Table 14. Paired-Samples T Test on the perceived importance to use higher-order thinking 131
Table 15. Mean score of the perceived importance to use lower-order thinking in the innovation class.....132
Table 16. Analysis of One-way ANCOVA on the comparison of question-answer behaviour.....152
Table 17. Analysis of One-way ANCOVA on the comparison of post question wait-time.....152
Table 18. Mean scores of post-question wait-time.....152
Table 19. Paired-Samples T Test on the perceptions of post-question wait-time in the innovation class.....153

Terminology

The following are the terminologies used in the current study.

L1: First language used by learners

L2: Second or foreign language learners study as a subject

Thinking skills: Skills refers to expertness, practical ability or facility in the processes of thinking, such as remembering, forming concepts, planning what to do and say, imaging situations, reasoning, solving problems, considering opinions, making decisions and judgments, and generating new perspectives

HOTS: Higher-Order Thinking Skills are analysis, synthesis and evaluation thinking skills described in Bloom's Taxonomy

Question: It refers to an utterance or a sentence that seeks information

Higher-order questions: Questions which require students to manipulate the information with higher-order thinking skills

Thinking tasks: Tasks have higher-order questions embedded within them.

HOTS approach: An approach which is used to develop students' Higher-Order Thinking Skills and speaking in an L2 social context. It has two steps. The teacher models the tasks to the students which focus on how to answer higher-order questions, interact and comment on others' opinions. Students then practice the thinking tasks in groups.

Socio-cognitive conflicts: Cognitive conflicts arising from a social context

Collaborative learning: An instruction method where students at various performance levels work together in a small group towards a common goal

ZPD: Zone of Proximal development, the place where students are unable to learn on their own, but can continue to learn when guided by significant others

Scaffolding: The temporary assistance which helps learners to accomplish a task beyond their actual level of ability

Critical thinking: a mode of thinking about any subject, content, or problem where the thinker improves the quality of his or her thinking by skilfully analysing, assessing, and reconstructing it; critical thinking can be categorized as a type of high cognitive thinking

High cognitive interactive talk: A tool of cognitive activity arising in a social context

High cognitive monologue: A tool of cognitive activity arising within an individual

CHAPTER 1 - EXPLORING THE RESEARCH CONTEXT

1.1 Introduction

Learning to speak in English has always been problematic for students in Taiwan. Chen (2002) found that they particularly fear oral-oriented activities in class. Among the four language learning skills, speaking is considered by many Taiwanese learners as the most difficult and yet essential skill to acquire (Hsu 2004). It is common for many learners to be unable to communicate freely in English even after learning English conversation for many years (SZE 1995; China Times e-paper 2009). To identify English (L2) speaking problems in Taiwan, relevant literature was reviewed and an exploratory study conducted to explore the current situation.

1.2 Outlining the issues

It is generally agreed that English speaking proficiency of most university students in Taiwan is low (Nunan 2003; China Times e-paper 2009). A survey conducted by Zhong (2001) showed that students at technology colleges throughout Taiwan believe that developing speaking ability is their greatest need. However, students' motivation to speak English in class is relatively low (Ho 2002) with a lack of learning and interest (Yang 1992). Moslehpour and Chou (2004) pointed out that students' low motivation to speak English is due to difficulties in learning and a high level of anxiety, which leads to an unwillingness to talk in class and limited learning outcomes. In particular, non-English majors are shown to have significantly higher English speaking anxiety than English majors (Liao 2008) in Taiwan. Yang (1992) and Shou (1995) found that the reasons for this low learning motivation in speaking classes are as follows:

a) The learning process focuses on correcting errors in form. This causes anxiety and nervousness from the students. Consequently, the motivation to learn to speak decreases.

b) Class sizes are too large, approximately 50 to 70. This reduces the opportunities for individual students to interact with the teacher.
c) The predominant disposition to learning is passive, with teachers dominating the instruction context, topic and speaking sequence. Instruction is teacher-centred instead of learner-centred.
d) Material is often irrelevant to the students' experience.
e) Few contexts similar to daily life for practicing English conversation are provided.

1.3 Exploratory study

An exploratory study was conducted in April 2007 to explore the current situation with regard to English speaking in Taiwanese university settings. The research question was: what is the predominant teaching relating to speaking in L2 classrooms? Data were collected through classroom observation and interviews. Observations were conducted in seven L2 classrooms in three universities located in northern, central and southern Taiwan. The researcher sat in the class and took notes, focusing on the classroom talk, including the amount teachers talked, the type of questions asked and students' responses.

Semi-structured interviews were conducted in Chinese with ten students and eight lecturers of English. Students were asked two questions: a) How did you feel when speaking English in class? b) Do you think the speaking opportunities provided in class are sufficient? While the teachers were required to respond to two questions: a) Can you describe a situation when students speak English? b) Can students think critically in class? Data were analysed by a process of content analysis (Holsti 1969).

1.3.1 Analysis and findings

The results gained from the classroom observations showed four features: a) teachers talked most of the time, b) teachers often asked display questions which are questions

teachers ask to check whether learners understand or remember something, c) students showed fear when nominated to answer a question, and d) students' utterances tended to be short, as noted in the following:

> '10 out of 12 classes, the teacher talked approximately 30 to 40 minutes in a 50-minute lesson. Occasionally, questions would be posed. Very few students volunteered to answer the question, therefore, the teacher needed to nominate students to answer it... Questions teachers asked were mainly display questions, such as asking the meaning of a sentence or a word and checking students' reading comprehension... Sometimes students were required to state their own reasons... When students were answering questions, more than half of them seemed embarrassed and anxious... Their utterances were mainly one to two sentences long.' (Observation note, April 2007)

The researcher also noticed an interesting phenomenon that when students elaborated on opinions, their utterances became longer. A lecturer who was a native speaker of English conducted a problem-solving task, which required students to think critically. During the group discussion, it was found that students tended to produce longer utterances stretching their limited English proficiency.

The findings obtained from the student interview data revealed two main problems encountered by students. First, six out of ten interviewees claimed that they did not have the courage to speak English because of a perceived inadequate vocabulary and poor knowledge of grammar, and this led to low motivation as stated by students, S3 and S6:

> S3: 'It is very difficult for me to speak in English because there are a lot of words I don't know and I'm afraid of making mistakes.'
>
> S6: 'I don't like to speak English because it's not fun and the teacher did not provide the relevant vocabulary and sentence patterns needed for discussion.'

Second, four out of ten interviewees, who were more confident in speaking English, pointed out that opportunities for speaking needed to be increased as claimed by S5:

S5: 'I wish we can have more chance to speak English in class.'

One observed example exemplifies this situation: a student replied to a teacher's question by saying 'I don't think so', and the teacher did not give the student a chance to elaborate his thoughts, instead, the teacher provided her own reasons to the class. Such behaviour could be seen to seriously disrupt students' access to language use and to thinking critically. The fact that teachers dominated the talk exemplifies the need for more speaking opportunities for the students.

Six out of eight teachers claimed that students' observed learning behaviours were affected by high anxiety and low motivation. As reported by the teachers, T4 and T6:

T4: 'Most of my students dare not to speak English in front of the class and their motivation seems rather low.'

T6: 'Freshmen do not want to speak English. They seem to have higher anxiety [towards English speaking] than those senior students.'

Five out of eight teachers reported a lack of risk taking ability as claimed by T8:

T8: 'My students seem to be afraid of making mistakes and they rarely try out the words they learned in different contexts.'

The disposition to take risks is defined in the developmental literature as engagement in behaviours that are associated with some probability of undesirable results (Furby and Beyth-Marom 1992). In this L2 context it meant that students were afraid of making mistakes and did not have the courage to try out the language learned in a new context.

Six out of eight teachers stated that students could not think critically and their reasoning ability was considered low, as reported by T2:

> T2: 'My students can't think critically. I once asked students to debate on whether college students should wear uniform and they couldn't find the reasons to argue or persuade their opponent group.'

1.3.2 Discussion

The observation data showed that the teachers talked most of the time. This is consistent with Walsh (2002) and Chaudron (1988) who found L2 teachers had the majority of classroom speech, and this remains one of the key features of teacher-centred classes where students do not receive enough opportunity to talk and communication is very limited (Legutke and Thomas 1991).

The classroom observations also revealed that teachers often asked display questions which is commonly found in the literature (Long and Sato 1983; White and Lightbown 1984; Cai 2003; Tan 2007). A possible explanation for teachers asking more display questions is that these questions are often asked as tests of the students' mastery related to particular points of language structure or vocabulary (Long and Sato 1983). This type of interaction pattern is described by Sinclair and Coulthard (1975, 1992) as 'Initiate Respond Evaluate' (I-R-E) or 'Initiate Respond Feedback' (I-R-F). In order to manage a large class the teacher needs to be in control of both the content and turn taking in the classroom.

Display questions used in L2 classrooms may promote effective transmission of linguistic knowledge. However, it needs to be recognized that such type of questions do not provide greater opportunities for students to talk; this as Thornbury (2000) argued, denies language learners access to what they most need in a language class: opportunities for authentic language use. All students needed to do to answer these

questions was recall knowledge. The cognitive effort involved is low. As Smith and Paterson (1998) and Coyle (2002) stressed, cognitively undemanding work like copying or repetition, does not enhance language learning, and such transactional 'scripts' do not work for the majority of learners. This finding implies that it is necessary to create opportunities for authentic language use.

The findings obtained from these data showed that students' utterances became longer when working on a problem solving task. It meant that when students used high cognitive thinking, they were likely to speak more. This finding is supported by research, for example, Cole and Williams (1973), who found a positive relationship between the use of high cognition and length of student response.

It was not surprising to learn from the interview data obtained from teachers that students had high anxiety and low motivation to speak English. In fact, studies on L2 anxiety show high levels to be associated with low motivation (Clément et al. 1994). This indicates that a context with low anxiety for L2 speaking is needed, and this should bring advantages for students' motivation to learn.

In brief, the findings from the exploratory study are consistent with the literature on L2 speaking problems in Taiwan. The exploratory study indicates that it is necessary to create more opportunities for students to talk and use authentic language. Students' learning motivation should be stimulated and the learning contexts provide a more relaxed environment where speaking anxiety can be reduced. Activating students' high cognitive thinking might be effective to enable students to talk more.

1.4 Context for the study

Having reviewed the problems related to L2 speaking, it is now important to explore the contexts of this thesis. This section introduces the role of the English language in Taiwanese Higher Education (HE).

English is recognized as an international language and is the most commonly studied foreign language in Taiwan, while Mandarin (Chinese) is the official language. English is studied as a required subject from elementary school to university. It is one of the main subjects considered at entry level for high school and university. In addition, English proficiency is often tested when applying for government jobs. Exposure to English use outside the language classroom is rather limited; however, additional English lessons outside the school are common and widely accessed (Tsai 2007).

One of the main aims of HE instruction in Taiwan is to enhance students' rational thinking skills and creativity (National Yunlin University of Science & Technology 2007; Ministry of Education of Taiwan 2001, 2007a, 2008). Rational thinking ability refers to the ability to analyse, compare, evaluate and make objective judgements. It is an important element for developing students' readiness for future work. Current education reform aims to reduce the use of the traditional learning style: rote learning by equipping students to think independently and proactively (Ministry of Education of Taiwan 2001).

Teaching thinking skills is an important part of the educational curriculum in many countries and a desirable goal in HE (Halx and Reybold 2006) and this is also reflected in second language (L2) education (Ayaduray and Jacobs 1997). Regarding the main aims of English language learning in Taiwanese HE, one is to enable students to communicate in English, express their opinions on social event or news which requires high cognitive thinking and use thinking skills such as analysing,

reasoning, evaluation, brain-storming, comparison, and problem solving while practicing the four skills (National Yunlin University of Science & Technology 2007).

Projects on excellence in HE instruction have placed an emphasis on developing students' English language proficiency due to 'internationalisation' (Ministry of Education of Taiwan 2007b, 2008). Universities have devoted much effort to enhancing students' English proficiency and learning motivation; changes include grouping students into various classes based on their English proficiency, offering extra English classes for low achievers, providing additional English conversation and English writing tutorials and requiring graduates to meet an English proficiency threshold (Tsai 2007) evaluated, for example, by GEPT (General English Proficiency Test). Seminars and conferences take place every year to offer English teachers opportunities to enhance their professional knowledge and skills (Ministry of Education of Taiwan 2008).

Having explored the aims and importance of English language learning placed by Ministry of Education in Taiwan, it is now important to explore common pedagogy in L2 classrooms: English lecturers and teaching methods. The majority of HE English language lecturers in Taiwan are Taiwanese (Liu 2005). Students who are non-English majors are predominantly taught by Taiwanese, whilst English majors often have the chance to be instructed by native English speakers. Moreover, Taiwanese lecturers mainly use bilingual instruction, Chinese and English, rather than English-only in English classrooms.

Two main teaching approaches are commonly used in English language teaching in HE, namely the Communicative Language Teaching (CLT) Approach and the Grammar-Translation Method (Liu 2005). The Ministry of Education's curriculum policy in Taiwan has firmly embraced the CLT approach introduced to L2 teaching in Taiwan

in the 1990s as the basis for pedagogic practice. In a CLT oriented classroom, students are expected to have various opportunities to communicate in the target language. However, it may be questioned as to whether this approach has been fully implemented in the way promised. As Coyle (2002) argues, this so-called 'communicative approach' in most L2 classrooms, with its prescribed syllabus topics based on transactional language, unintentionally promotes a reactive rather interactive role for learners. The finding of the exploratory study is an example of this. Students can be de-motivated and impeded by such pedagogy. A lack of teacher training also indicates CLT has not been fully put into operation in Taiwan (Nunan 2003).

In a class using the Grammar-Translation Method, lecturers mainly use Chinese to explain the grammar and meaning of texts. The Audio-Lingual Method is often used to train students' listening comprehension and speaking skills through drills. Students may be required to practice dialogues written in the textbook led by the lecturer, or sometimes in pairs or in small groups to develop their speaking fluency. However, some English classes apply both CLT Approach and Grammar-Translation Method (Liu 2005). The Grammar-Translation Method may be employed when working on reading comprehension, while the CLT Approach is used for listening and speaking skills.

These teaching approaches mainly focus on the gain of linguistic competence and do not offer many opportunities for students to think critically and independently and to speak authentically; students often sit quietly to listen, respond passively to those comprehension-checked questions and practice speaking with mainly written dialogues. These approaches arguably require a large amount of low cognitive thinking, for example, knowledge, comprehension and application, and to a certain extent they do help L2 learning. However, a consequence of these types of teaching is that students' English speaking proficiency generally is low, especially in

communication (Nunan 2003; China Times e-paper 2009) and this can be reflected by students' attitude towards English language learning: high anxiety and low motivation as discussed in Section 1.1. English courses at many Taiwanese universities seem to have not been successful enough to motivate students to learn (Huang 1998).

This trend is exacerbated by a general Taiwanese students' learning style which tends to be passive. In many Chinese classrooms teachers tend to dominate the lesson while the students habitually just listen and respond passively (Holliday 1994; Yang 1992; Shou 1995). The exploratory study corroborated this view. Such classroom discourse has been noted as different to that seen in western classrooms, which tend to be filled with talk and noise (Nora 1997; Sato 1982). This could be magnified by research (Holliday 1994; Liu 1998) that shows Asian teachers are seen not as facilitators but as knowledge deliverers. Such rote learning is what the educational goal aims to reduce by promoting active thinking in class (Ministry of Education of Taiwan 2001).

In conclusion, HE in Taiwan has highlighted the importance of thinking and English language learning, in particular the skill of speaking. However, the implementation of English language instruction does not seem to carry out the aim of HE: enhancing students' high cognitive thinking ability. There appears to be a gap in L2 classrooms.

One issue which needs to be considered if high cognitive thinking is to be infused into L2 learning is whether students with a passive learning style are able and willing to manage active learning that requires students to think independently and critically. One way to promote active learning is through introducing high cognitive thinking. Regarding this, Fisher (1998) has postulated that thinking can be developed through training. Littlewood (2000) also stated that although Asian students adopt passive roles, this does not appear to reflect the roles they would like to have in class. They

adopt passive classroom attitudes as a consequence of the educational contexts they have experienced rather than any inherent dispositions of the students themselves. These statements imply that students may become active learners when the pedagogy used in L2 classrooms allows them to think independently, critically and creatively. With such a chance, this study aims to put high cognitive thinking into operation in L2 learning using the HOTS approach. The process by which thinking and language learning can work together to enhance students' performance will be developed further in Chapter 2. The following section will introduce the HOTS approach used in this study.

1.5 The HOTS approach

The HOTS approach was the innovation conducted in this study to develop students' HOTS and speaking in an L2 social context. In this study, HOTS are analysis, synthesize and evaluation described in Bloom's taxonomy (1956) (for more details, see Section 2.4). This approach had two steps. In a teacher-led setting the innovator modelled the tasks to the students which focused on how to answer higher-order questions, interact and comment on each other's opinions. Students then practiced the thinking tasks in groups as part of their L2 learning. These two steps were conducted weekly for 12 weeks.

This approach was different from the CLT approach and Audio-Lingual Method in terms of a practice of speaking. It required students to activate their schemata and use their higher-order thinking proactively in a social context to express their thoughts and comment on others' ideas. This can be more cognitive demanding and require more efforts than the traditional methods used. The following introduces further details of a design of the thinking tasks, the lesson plan and the implementation of this innovation.

1.5.1 Design of thinking tasks

Thinking tasks are tasks which have higher-order questions embedded within them to require students to actively use their higher-order thinking. Question in this study was defined as an utterance or a sentence that seeks information, thus, it can be a question or a statement in the tasks. Different thinking tasks contain different linguistic and cognitive demands, and can be placed in the resulting quadrants, as shown in Figure 1. These two aspects need to be taken into account while designing the tasks. To design cognitively challenging tasks, students' familiarity with the topic is essential. Similarly, the task can be linguistically demanding, yet manageable to students. Therefore, to ensure the accessibility of language use, links to help students to identify related sentence patterns and vocabulary should be provided.

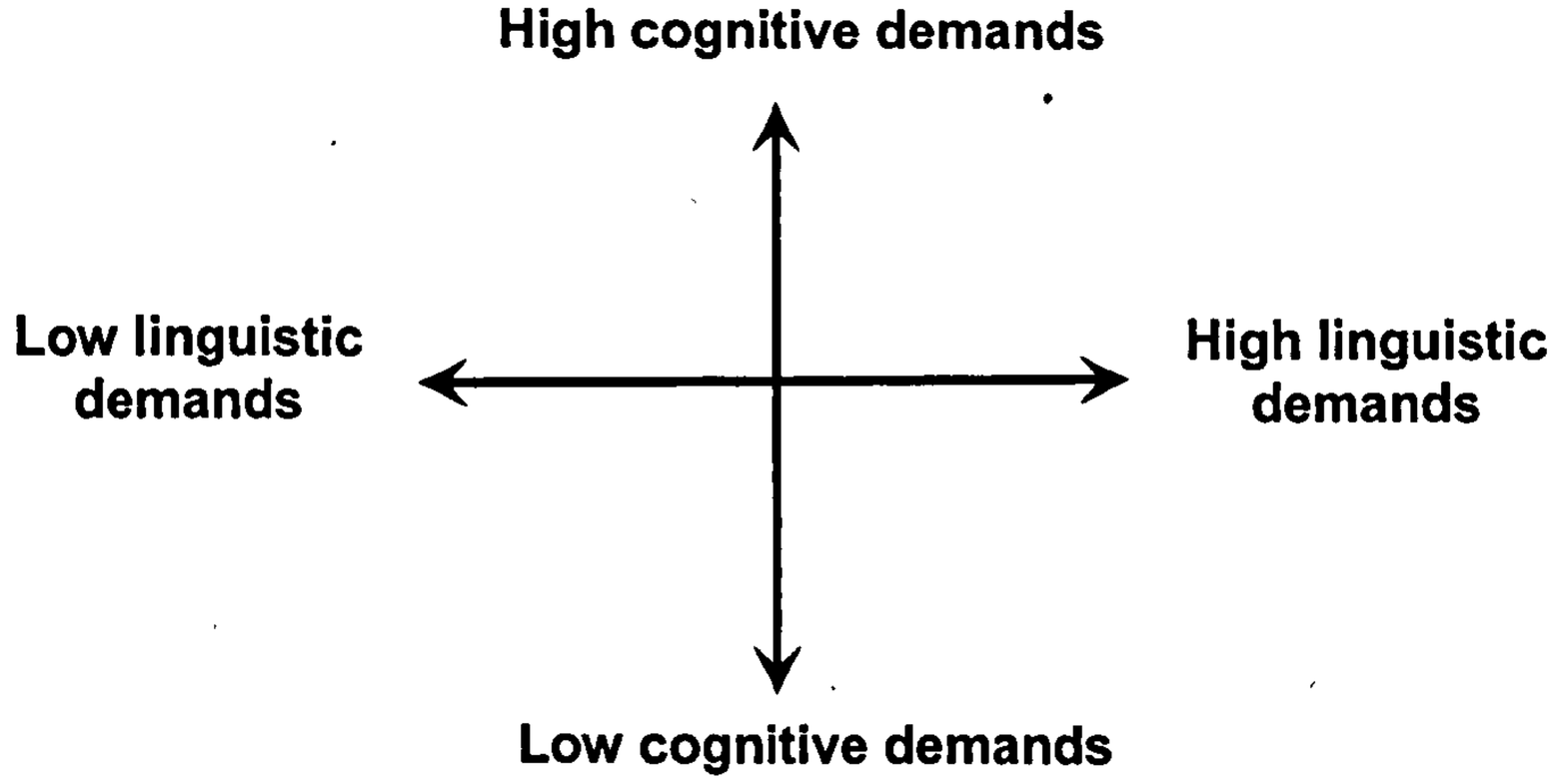


Figure 1. Quadrants of linguistic and cognitive demands in tasks

(Adapted from Coyle 2002)

Four types of thinking tasks, 5Ws, Odd One Out, Make Up A Story and Guess What I Say, were chosen in this study because each task provided opportunities to exercise a particular skill of higher-order thinking and with these four tasks all higher-order thinking skills were fully operated. The rationale to apply *four* tasks is to prevent fatigue towards the innovation. It needs to be recognized that there are other thinking activities which are available for teaching thinking (for other thinking tasks, see Lake and Needham 1995; Leat 1998; Higgins and Baumfield 2001; Nichols and Kinninment

2001; Lin and Macky 2004; Butterworth and O'Connor 2005). Tasks designed in this study aimed to activate students' schemata to develop thinking through the use of L2 speaking. Topics used were all related to the content of the textbook and life, e.g. travelling experience, the ideal mate, characters, travelling packages, pets and so on.

5Ws was adopted from Butterworth and O'Conner (2005) and it was used in the present study to activate students' higher-order thinking, particularly at the thinking level of analysis and evaluation. It mainly asked higher-order questions like 'why', 'how', 'which is the best', etc. An example of 5Ws task is as following:

> Skilful decision making
> Options: What could Hamlet have done?
> Based on the option taken, talk about:
> a) What would happen if he took this option?
> b) Why do you think that this consequence could have occurred?
> c) How important are the consequences? Why?
> (Butterworth and O'Conner 2005, p. 18)

This 5Ws task required students to actively use their thinking level at analysis synthesis and evaluation. This type of task was further enhanced in this study by creating a social context for discussion. In order to create a context for interaction, 5Ws was designed to have students reached a consensus in a group; students needed to reason and argue with each other.

Odd One Out is an activity that develops the skills of classification (Leat 1998). It required students to utilize the thinking level at analysis to categorize the three or four objects and identify which one is different from the other two and what the other two have in common. An example of this task is shown in Figure 2.

Try to decide which number from each set is the Odd One Out. Underline this word in your book and explain why it is the odd one out and what the other two have in common.

Set A	2	13	3
Set B	4	15	6
Set C	8	27	31
Set D	22	10	25

(Leat 1998, p.13)

Figure 2. Odd One Out

Make Up A Story was developed by the researcher and aimed to develop students' creative thinking. It requires learners to create a story based on four pictures provided. The pictures provided in this study were related to travelling, which had a story behind it. Therefore, after students had created their own story, the researcher shared the real life story with them. An example of this task is shown in Figure 3.

1. **NT$200,000**	2. **Italy**	3. **Tour guide**	4. **motorcyclist**
 NT$200,000	 Italy	 Tour guide	 Motorcyclist

Figure 3. Make Up A Story

Guess What I Say developed by the researcher required students to figure out the answer with analytical thinking. One student, for example, explains the meaning of a Chinese idiom in English without mentioning the key words, and the rest, based on the clues given, infer what the idiom is. Thinking required for the speaking is comprehension and application, while it required listeners to analyse the information given. Figure 4 is an example of this.

Chinese idioms:
食言而肥
騎虎難下

Figure 4. Guess What I Say

1.5.2 Lesson plan

Each type of task was designed with three different topics, totalling 12 tasks in which links to the related sentences and vocabulary were provided (for the thinking tasks used in this innovation, see Appendix 1). One task was used per week with a cycle of four types of tasks. How these tasks fitted into the weekly schedule will be presented below. Table 1 demonstrates an example of weekly lesson plan, which was a 100-minute lesson, and this weekly lesson plan was repeated for 12 weeks. For the first 50 minutes the class focused on the textbook and students learned the target vocabulary, sentence patterns and the text. After 10 minutes break the thinking task was then conducted. The process included modelling, managing the activity and de-briefing. The modelling was firstly conducted for 15 minutes: the innovator modelled the tasks to the students. The modelled tasks focused on how to answer higher-order questions and comment on others' views. Sometimes, the teacher would demonstrate how students could conduct that particular week's task, for example, Odd One Out, Make Up A Story or Guess What I Say. Students then spent 20 minutes to work with the thinking task and model what they learned in a group discussion forum, and the innovator worked as a facilitator at this stage and mediated critical questions discussed. Finally, the de-briefing stage allowed each group to elaborate their opinions in a class discussion setting and the students could comment on each group's ideas.

Table 1. An example of weekly lesson plan

Chapter 9 Countries

Learning focus: to learn the country names, cities, famous international monuments and sights
Activity: Lecturing, class discussion, group discussion,
Prior knowledge: Country names, geographic location, basic knowledge about the countries
Timing: Total: 100 minutes The first half of the lesson: Looking at the ***textbook*** 50 mins: Lecturing, introducing vocabulary, text ----------10 minutes break---------- The second half of the lesson: Conducting ***the thinking task*** 15 mins: **Modelling**- introducing the task, to activate students' schemata, Modelling to students to answer questions and interact in a teacher led setting 20 mins: **Managing the activity -** thinking task (group discussion) Task: 5Ws, thinking skills: analysis, evaluation 15 mins: **De-briefing** (in a teacher-led setting) (Procedure in conducting the task was adapted from Lin and Mackay (2004)
Learning Objectives: • to enhance learners' ability in talking about different countries, cities, famous international monuments and sights • to enhance learners' thinking skills: comprehension, application, analysis, creative, evaluation • to enhance learners' English speaking ability

1.5.3 Implementation

The innovation was implemented by the researcher who had seven years of English as a foreign language teaching experience. The formal English lecturer, who had ten years teaching experience, chose not to participate in the innovation. At the planning stage of the research, the formal lecturer had agreed to conduct the innovation following a lesson plan designed by the researcher. Before the semester started, the researcher had spent several days with her explaining what thinking skills and higher-order questions were about and how to conduct the higher-order questions and thinking tasks in class. The lecturer mentioned that she had encountered the concept of higher-order thinking before. However, after the pre-test was done with the comparison class she discovered that she did not have the confidence to conduct the innovation. Therefore, she suggested that the researcher should take over the

innovation class and she could work with the comparison class. Thus, the 12-week innovation was conducted by the researcher. After the innovation, the lecturer took over the innovation class following her own schedule with her own pedagogy. The impact of a change of innovator will be discussed in Chapter 6.

The innovation class received the innovation of the HOTS approach in both a teacher-led setting and group discussion forum. The comparison class did not receive it and they continued to practice speaking in both a teacher-led setting and small group discussion based on the speaking activities provided in the textbook (Yeldham 2001, for an example of a discussion exercise extracted from the textbook, see Section 7 Discussion in Appendix 2).

An example of modelling in a teacher-led setting is described, taking the task of 5Ws as an example. First of all, the researcher explained what the task was about and what students needed to do. Students' schemata were activated by the related questions concerning the topic. Then a higher-order question was posed, a wait of at least five seconds was given (for detail, see Section 2.7), but in practice the researcher found it insufficient. Wait-time was given until the students came up with answers, which was more than 30 seconds at the beginning of the innovation. With the response given, post-response wait-time was provided. The researcher, then, would comment on the student's opinion by saying 'I disagree/agree with you because...' or 'In my opinion, it is better to choose... because...' (Lipman et al. 1980). Following this, other students would be invited to make further comments. Sometimes, probing was used to further explore students' thoughts, reasons or related information by asking e.g. 'Why do you think it's better for her?' or 'Tell me more about it' etc. In this way, the researcher modelled how to express thoughts, comment on others' ideas and further probe for related information needed. After the modelling, students then modelled what they had learned in small group work. During group discussion, the researcher played a role as

a conversation facilitator. The researcher listened to students' conversation and asked occasional questions to encourage elaboration which allowed learners to articulate their thoughts more clearly and logically. After group discussion, class discussion was conducted where each group talked about their opinions and reasons with the others responding accordingly.

1.6 Purpose of the study and research questions

The purpose of this study was to promote Higher-Order Thinking Skills (HOTS) in a Taiwanese university L2 classroom with the aim of enhancing students' speaking and thinking. It aimed to evaluate the practicability of this approach by examining the impact on students' speaking and thinking; while students' attitudes were also taken into account. A better understanding of the extent to which this approach impacts on students' learning and achievement was explored.

This study comprised one main research question with three sub-questions which were felt to meet the key issues identified through literature review and exploration of the context. The main question was:

Is it practicable to use the HOTS approach to develop speaking ability and foster higher-order thinking skills in a Taiwanese university L2 classroom?

The sub-questions were:

1) How does the introduction of the HOTS approach impact on students' oral performance?
2) How does the introduction of the HOTS approach impact on students' observed cognitive behaviours?
3) What are students' attitudes towards and perceptions of the impact of the HOTS approach?

1.7 Value of this study

Having stated the purpose of this study and the research questions, it is important to explore briefly the value of this study. This research explored the use of the HOTS approach as an innovation and examined its effects on L2 speaking and cognitive performance and students' attitudes in Taiwan. Previous experiments using higher-order questions in L2 classrooms have been mainly conducted in western countries, for example, American and Spain, and former British colonies like Singapore and Malaysia. This study critically looked at the practicability of the HOTS approach in the Taiwanese context, where class size tend to be large, students' learning style are predominantly passive, and the use of L2 outside classrooms is relatively limited. The findings of this study are expected to benefit researchers in this field, language teachers and teacher trainers with a better understanding of how a HOTS approach can support L2 learning in a HE setting. Thus, this study aimed to tackle the dual aims of the current Taiwanese Government to increase thinking and create better English language speakers for participation in the global economy (Ministry of Education of Taiwan 2008). Chapter 2 will extend the literature base to explore how thinking and language learning can work together simultaneously by exploring the nature of thinking skills and the extent to which such skills help students' L2 speaking and cognitive performance.

CHAPTER 2 - LITERATURE REVIEW

2.1 Introduction

Having explored L2 speaking problems and identified the importance of thinking skills in HE and second language teaching, this chapter reviews the literature on how thinking skills enhance student oracy and cognitive function. It first examines learning theories that focus on cognitive development and second language acquisition, drawing a general picture of how cognition is cultivated and how L2 speaking, in particular, is acquired. It then turns to review thinking skills programmes and their effectiveness so as to inform the potential effects in L2 classrooms. This is followed by a discussion of teaching and learning associated with thinking, identifying questioning as an appropriate means to scaffold thinking and speaking by providing the oral interaction and space to think. To formulate questions, Bloom's taxonomy is adopted as a thinking framework because of its systematic classification of cognition which is easy to apply. It further examines the effectiveness of higher-order questions and strategies related to questioning and critically reviews the questioning behaviour. These inform the design of theoretical conceptual framework and help to identify the research gap and form the research questions.

2.2 Theories of learning

This section looks at the learning theory, social constructivism. It then moves on to articulate the theories of second language acquisition: Swain's output hypothesis, Long's interaction hypothesis and Gass's model of second language acquisition.

2.2.1 Social constructivism

Vygotsky (1978) claimed that a mechanism for individual cognitive development is social interaction. An individual gains understanding by constructing new knowledge or transforming old knowledge into new, and the process is facilitated through verbal

interaction. Learning occurs in a social context with the scaffolding (Wood et al. 1976) of other people in the Zone of Proximal Development (ZPD) (Vygotsky 1978). Such a learning process through social interaction influences both the development of language and intelligence. This learning theory can be linked to L2 learning, for example, Swain's output hypothesis (Section 2.2.2) where language improves when it serves as a cognitive tool in a social context or Long's interaction hypothesis (Section 2.2.2) where speaking is enhanced through meaning negotiation. These two L2 learning theories both emphasize the importance of social interaction on language learning. It is worth noting that in the ZPD, assistance not only can be provided by the more skilful learners, but also non-experts in L2 learning (e.g., Donato 1994; Ohta 1995; Swain and Lapkin 1998).

Vygotsky (1978) divided language used in a social context into two areas. First, language arises initially as a means of communication between people and their environments. Second, language may be used as a tool to mediate the self as well as one's environment. In a second language classroom, language is the target to be learned. Thus, the target language used in this study functioned as a communication tool, medium for cognitive activity and the learning objective.

A social constructivist approach is one where the learning is constructed jointly through social interaction (Vygotsky 1978). It allows the occurrence of socio-cognitive conflicts which contributes greatly to cognitive development. During the social interaction, different perceptions, ranging from simply having more or less schemata to holding completely contradictory perspectives, arise and are readjusted. Students are forced to externalize their thoughts, making their ideas explicit to themselves and to others. Following this, the continuous commenting, justifying and arguing provides them with opportunities to discover and fill the gaps in their knowledge structures,

correct misunderstandings, recognize and resolve discrepancies in information, and then readjust conflicting view points (Mugny and Doise 1978).

Cognitive conflicts arising in a social context like this have greater cognitive benefit for an individual than the conflict of ideas that an individual might experience alone (Doise and Mugny 1979) and through such a process knowledge can be restructured (Doise and Mugny 1979; Bruner 1986). To link the ideas of socio-cognitive conflicts with this study, the HOTS approach was designed to allow students to express their opinions in a social context, to justify other's ideas, and to argue and reason with each other in order to reach a consensus in group discussion. This is where students proactively used their higher-order thinking. In the process of making their own thoughts explicit to themselves and others, they needed to try out their L2 linguistic hypotheses and the comprehensibility and this facilitated the enhancement of their language ability. During the interaction, opportunities could occur for students to scaffold each other to think and use the target language.

A common feature of a social constructivist approach is collaborative learning (Webb and Palincsar 1996; Gokhale 1995). Proponents of collaborative learning (for example, Johnson and Johnson 1986; Gokhale 1995) claim that the active exchange of ideas within small groups not only increases interest among the participants but also promotes higher-order thinking. The language use in this verbal interaction can be seen as authentic where students are facing uncertainty. With the HOTS approach students were required to work together to achieve a common goal through verbal interaction. It is such verbal interaction or active exchange of ideas that creates more opportunities for students to think and to talk and thus facilitate thinking (Long and Porter 1985; Cam 1995; Mercer 2000; McGregor 2007) and L2 development (Seliger 1977; Zhou 1991; Gass and Varonis 1994; Mackey 1999).

The types of talk formed in this kind of collaborative learning are described by several researchers. Mercer (2000), for example, has identified three types of talk in L1 classrooms: disputational talk which is characterised by disagreement and individualised decision-making, cumulative talk where speakers build positively but uncritically on what the other has said and is characterised by repetition, confirmations and elaboration and exploratory talk in which students engage critically but constructively with each other's ideas. *Exploratory talk* can be more effective for learning because it embodies the principle of constructive criticism (Mercer 1995). *Dialogic talk*, described by Alexander (2003), is collective, reciprocal, supportive, cumulative, purposeful and productive where questions, answers and feedback progressively build into coherent and expanding chains of enquiry and understanding. In L2 collaborative learning Swain (2000) and Swain and Lapkin (2002) identify the talk as *collaborative dialogue* where speakers are engaged in solving linguistic problems and building linguistic knowledge. Exploratory talk, dialogic talk and collaborative dialogue share a common feature: providing the opportunity for students to extend their talk and thinking. The HOTS approach followed this kind of argument and aimed to provide more opportunities for speaking and thinking by requiring students to reach a consensus in group work with topics relating to daily life rather than focusing on linguistic issues.

A social constructivist approach has also been shown to facilitate the learning of low achievers. Vygotsky (1978) and Watson (2000, 2001) believed that social interaction could promote low achievers' cognitive development and enhance their learning. Research (e.g. Quicke and Winter 1994; Powell and Makin 1994) showed how the use of a social constructivist approach could support low achievers; they became more able to share ideas, express opinions, and talk about their learning. This feature could be very important in the HOTS approach in terms of helping students with low speaking proficiency. It is expected that through the speaking and thinking

opportunities created low achievers can gain more chances to try out their language use.

The advantages of a social constructivist pedagogy, as stated in Watson (2000), can be seen as the promotion of learning experience in which learning is active, thinking is built on and extended, the awareness of one's own learning is assisted, transfer is facilitated, students are in control of their own learning, and confidence and self-esteem are raised. These advantages can be very useful in L2 learning. For example, students can control the topic discussed and dominate the talk and thus active learning is promoted which is one main educational goal. Besides, the transfer of L2 outside the school is the main aim of L2 learning: the core issue is how good students can apply the language learned to other domains like work places.

2.2.2 Theories of second language acquisition

Verbal interaction is one of the main features of social constructivism. Swain (1985) linked this process with L2 acquisition and stated that learners need the opportunity to talk in order to develop language competence. In Swain's (1985) output hypothesis,

> '[comprehensible output provides] opportunities for contextualized, meaningful use, to test out hypotheses about the target language, and to move the learner from a purely semantic analysis of the language to a syntactic analysis of it.' (Swain 1985, p. 252)

Swain argued that language production may facilitate second language learning in ways that complement the role of input in L2 learning, specifically by encouraging learners to notice linguistic forms in the input, test their hypotheses about how the target language works, and to use language to reflect on language use (Swain 1995, 1998). Since originally formulating this hypothesis, Swain (2000, 2005, 2006) has moved from viewing output solely within an information-processing framework of

learning to viewing output from a sociocultural perspective. An essential difference is that the latter considers output not only as a message to be conveyed, but also as a tool for cognitive activity.

From a sociocultural theory perspective, Swain argues that language serves as a cognitive tool that assists in the learning process. By producing language, learners can construct and co-construct L2 knowledge through interactions with interlocutors and by reflecting on their own language use. With the HOTS approach students have chances to express their own opinions and comment on others' ideas and these opportunities allow students to try out their language hypotheses and comprehensibility. The output produced with this approach mainly focuses on conveying meanings to others and this arguably is the most effective way to learn a language (Richards and Rodgers 2001; Williams 1998).

Swain's output hypothesis mainly focuses on the output produced facilitating the learning of speaking by trying out linguistic hypotheses and comprehensibility, but this is insufficient in terms of looking at the how speaking competence can be enhanced through verbal interaction in a social context. Long's interaction hypothesis helps to explain this. In Long's interaction hypothesis (1981, 1983, 1985, 1996) the negative feedback from meaning negotiation may induce the noticing of some forms.

> 'it is proposed that environmental contributions to acquisition are mediated by selective attention and the learner's developing L2 processing capacity...negative feedback obtained in negotiation work or elsewhere may be facilitative of SL development.' (Long 1996, p. 414)

In the HOTS approach, the exchange of ideas provides the opportunities for negotiation of meaning. Through negotiating meaning, students may ask for clarification or confirmation and this allows the speaker to have opportunities to

understand their produced output that is incomprehensiable and thus revisit his/her linguistic hypothesis, modify and try it out again. This type of interaction mainly requires lower-order thinking like knowledge, comprehension or application (Section 2.4.1) and it, arguably, does not create greater opportunities for talk. Long's interaction hypothesis, however, does not explain what type of interaction can create more opportunities for students to try out the linguistic hypothesis. This aspect will be explored in this study with the HOTS approach. Swain and Long have looked at parts of second language acquisition. To gain an overall picture of how second language is acquired from input to output, Gass's model of second language acquisition is examined.

Gass (1997) proposed a more comprehensive model with regard to second language acquisition, suggesting that the learner passes through five phases from noticing L2 input to producing L2 output during the dynamic process of acquisition: 1) apperceived input, 2) comprehended input, 3) intake, 4) integration, and 5) output. Apperceived input refers to 'that bit of language that is noticed in some way by the learner because of some particular recognizable features' (ibid., p. 4) and these features are related to the learner's prior knowledge. Not all noticed input becomes comprehended input. Whether comprehended input becomes intake is partly decided by which level of analysis is reached. In Gass's model, intake refers to 'the mental activity that takes learners from the input to their interlanguage grammars' (ibid., p. 137). Intake can either be integrated into the learner's implicit knowledge system or put into storage. Finally, the produced output can function as feedback for testing a hypothesis and as moving from semantic to syntactic processing. This model concludes that 'second language is shaped by the input one receives and by the interactions in which one engages' (ibid., p. 161).

In Gass's model, it is pointed out that some variables like frequency, motivation, attitudes, negotiation of meaning are influential in terms of a transfer of comprehensible input to the interlanguage system. It can be argued that with the HOTS approach a discussion around a topic at a time can increase the frequency of encountering certain vocabulary or linguistic structures mentioned by the speaker or the others; this is unlike the input received in a teacher-led setting or in an activity where interaction is not carried out actively or enthusiastically. It is very important to highlight the opportunities to use the target language, which can be created by the HOTS approach due to a demand of exchange of ideas. Topics used based on students' life experience arguably have greater chances to motivate students to talk and use the language. In order to follow the conversation or get a point across, negotiation of meaning can occur, as stated in Long's interaction hypothesis. In particular, spoken language can be used as a cognitive tool in the HOTS approach and this would be further conducive to the L2 learning process, as argued by Swain (2000). Such idea links with Allwright's (1979) emphasis on the importance of meaningful interaction in L2 learning rather than rote repetition or study of language as object. These aspects discussed are all the strengths of the HOTS approach which can facilitate L2 learning and the extent to which this approach can provide optimal conditions for the language learning is explored in the following section.

2.2.3 Summary: exploring optimal conditions for L2 learning

The optimal conditions for classroom language learning based on second language acquisition theories, as proposed by Ellis (1990), are that learners must have the need to use the target language, and the opportunities to initiate and control topics. Learners are supposed to be able to participate in planned and unplanned discourse at any time the need occurs. The teacher, on the other hand, should provide scaffolds for learners to produce language structures, which are beyond their current level. This kind of optimal interaction is more likely to happen in meaning-focused instruction than

in form-focused instruction. The HOTS approach focuses on the meaning-focused interaction where students can dominate the talk and initiate and control topics; this type of discourse is unplanned with simultaneous interaction and has characteristics of unpredictability and uncertainty. These learning conditions arguably offer most advantages for the learning of L2 speaking. In addition to L2 learning, one main aim of HE in Taiwan is to enhance students' thinking ability so as to decrease rote learning. Therefore, it is now necessary to turn to examine thinking skills.

2.3 Thinking skills

This study aimed to infuse thinking skills into L2 learning. Having explored the overall learning theories of cognitive and language development, it is now necessary to explore what thinking skills are, how they work in L1 classrooms and their status quo in L2 classrooms. In order to inform the theoretical conceptual framework of teaching thinking, the most suitable means for teaching and learning of thinking is explored.

2.3.1 What are thinking skills?

There is no universal agreement as to the precise meaning of the term 'thinking skills'; however, different definitions generally share common ground which refers to the human capacity to think in conscious ways to achieve certain purpose. This study adopts Moseley et al.'s (2004) definition of thinking skills which refers to expertness, practical ability or facility in the processes of thinking, such as remembering, forming concepts, planning what to do and say, imaging situations, reasoning, solving problems, considering opinions, making decisions and judgments, and generating new perspectives. In the process of thinking with the HOTS approach students are not only using Higher-Order Thinking Skills, but also recalling knowledge, planning what to say in terms of both the content and language use, forming concepts, reasoning, arguing, solving problems encountered, considering others' opinions and critically making judgements and commenting on it, finally reaching a consensus by making decisions.

The thinking process occurring in this approach corroborates with Moseley et al. (2004), therefore the definition was adopted.

The study aimed to enhance students' thinking through the treatment of HOTS. The thinking process occurring in the HOTS approach has much broader effect other than the development of HOTS itself like the improvement of reasoning skills, problem solving, consider others' opinions (see Chapter 5 for detail). Therefore the title of this thesis refers to infusing *thinking skills* rather than *HOTS*.

2.3.2 Thinking skills programmes in L1 classrooms

Thinking skills programs or thinking skills approaches are pedagogy which aims to activate students' thinking skills and enhance learning. McGuinness (1999) states, that the teaching of thinking is a general thrust in educational reform which emphasises the quality of thinking processes and thinking skills as a means to raise educational standards and to prepare learners for lifelong learning. Thinking skills programmes have been extensively used around the world, and have begun to be an important topic and a desirable goal in HE (Halx and Reybold 2006). Halpern (1999) stated that developing students' critical thinking abilities is a desirable if not the principal goal of HE in America. This is also one of the main aims for HE in Taiwan (Section 1.4). Education aims not only to impart knowledge, but also to enable students to probe questions and facts and to use thinking skills to solve problems (Bruner 1972).

There are many thinking skills programmes which have shown a positive effect on improving learners' performance in cognitive and curriculum tests (Lipman et al. 1980; Feuerstein et al. 1979; De Bono 2000; Eisenman and Payne 1997; Thinking Skills Review Group 2004; Higgins et al. 2005). In addition, Moseley et al. (2004) found that thinking skills approaches help learners to plan, describe and evaluate their thinking and learning. Cognitive enhancement resulting from thinking skills training could be

maintained in L1 classrooms. For example, Yang et al. (2005) found that students maintained their critical thinking skills after exposure to and modelling of Socratic questioning in an asynchronous discussion forum. Furthermore, Zohar's (1996) study also showed that students were able to retain the reasoning skills they acquired across time.

Several empirical studies completed in L1 classrooms have shown that argument-based reasoning skills can be transferred to other domains, for instance, Hunter-Grudin's (1985), Zohar's (1996) and Zohar and Nemet's (2002). In addition, Miri et al. (2007) found that with the promotion of *HOTS* in science classes, students were capable of transferring higher-order thinking across domains. Transfer is the ability to extract a particular skill from its original context and apply it to a novel situation (Perkins and Salomon 1988). Cognitive transfer has been documented as one of the ultimate goals when teaching students to think (Zohar and Dori 2003; Ritchhart and Perkins 2005; Reece 2007). The literature on thinking skills in L1 classrooms can function as a predictor for how thinking skills might operate in L2 classrooms and their impact on student performance.

2.3.3 Thinking skills in L2 classrooms

One main aim of learning English in Taiwanese HE is to practice the four skills with thinking skills, as discussed in Section 1.4, so as to enable students to talk critically while expressing their views. The literature showed that thinking skills programmes are mainly used in L1 classrooms. Infusing thinking skills into L2 learning has been initiated in some classrooms, but it is still uncommon and considered a relatively unresearched area (Lin et al. 2010; Waters 2006). High cognitive thinking has become a component of L2 reading texts and composition (Day 2002) and some textbooks (e.g. Tanka and Most 2007; Wegmann and Knezevic 2007) have adopted it into

activities designed as written or spoken exercises. Yet, the extent to which teachers of English language implement this kind of teaching material remains unknown.

It would appear a logical link for thinking skills to be introduced into L2 classrooms and due to the important role they play in HE development then the added benefit would appear to be worthy of investment in this sector. Van Duzer and Florez (1999) advocate that language learning should go beyond the learning of basic literacy skills because with the increased focus on information in today's world, the way people process information has become more important than the information itself which requires high cognitive thinking. Ustunluoglu (2004) further points out that language classes are particularly appropriate for teaching thinking skills owing to the richness of material and the interactive approaches used.

There are those however who have presented arguments against using a critical thinking pedagogy in L2 classrooms. Atkinson (1997), for example, argues that such approaches are inappropriate. He argues,

> '[we] should give TESOL educators pause for thought, and pause long enough carefully and critically on the notion of critical thinking' (ibid., p. 89).

Reasons given for this argument are concerned with critical thinking as being a social practice; the difficulty in teaching thinking to L2 speakers are mainly articulated as cultural problems and the complexity of transferring thinking skills beyond the narrow contexts of instruction. Atkinson (1997) points out that critical thinking is much less practiced in Chinese and Japanese cultures where group solidarity is more highly regarded than individual self-expression in society, and memorization and recitation are promoted as major learning strategies in schools. Moreover, Fox (1994), Ramanathan and Kaplan (1996a) have claimed that L2 learners are not ready for critical thinking courses and it is difficult for non-native graduate students to learn the

notion of critical thinking in academic writing. Scollon (1991), for example, found that Taiwanese university students had difficulty in comprehending the U.S. style process writing instruction which required them to think critically with a rational mind, much less to write from. These researchers all postulate that it is difficult or impossible for non-native English-speaking students, in particular Asian learners, to think critically because of their collective and hierarchical cultural backgrounds where critical thinking is not a common social practice and especially students rarely challenge what they learned from the teacher.

There are, however, advocators who consider critical thinking suitable for L2 classrooms (Gieve 1998; Raimes and Zamel 1997; Spack 1997; Zamel 1997). Day (2002) discovered that non-native students from Taiwan, China, Korea, and Japan studying at a university in America had no difficulty in engaging in critical thinking. But this was out of their home countries. Although students in Taiwan still tend to gain most knowledge from their teachers and not challenge what they have learned in classes, Lin's (2005) intervention study demonstrated that undergraduate students can be motivated to learn English through using critical thinking in the L2 classroom. This evidence, together with the discussion of the possibility to convert the passive learners to active learners (Section 1.4), indicates a greater possibility of applying thinking skills to L2 learning. So now it is necessary to explore how thinking can be taught and learned in L2 classrooms.

2.3.4 The teaching and learning of thinking

Thinking skills are often taught in classrooms. They are practical skills that focus on knowing 'how' rather than knowing 'what' and can be developed through experience, education, and training (Fisher 1998). According to Sousa (2001), teachers do not teach the brain to think, but thinking skills certainly can be taught at all levels to improve learners' achievement. Learners can be equipped with the skills to search out

meaning and impose structure, to deal systematically yet flexibly with novel problems and situations, and to communicate effectively.

Thinking skills can be taught as an isolated subject or through integration with major subjects, for instance, mathematics, English, etc. Thinking skills can be seen as learning strategy, in which cognitive abilities such as memorization, comprehension, deductive reasoning, inference and creativity are used (Oxford 1990; Wenden 1991). Wenden (1997) suggests that the integration of strategy instruction into regular language instruction may be the most effective approach, as did this was reached by Ayaduray and Jacobs (1997). Jones et al. (1987), on the other hand, point out that, when taught in isolation, knowledge and skills may not transfer across the curriculum, especially with less proficient students.

Paran (2003) believes that incorporating elements of high cognitive thinking into material encourages learners to question texts and add linguistic value to the textbook and classroom. This notion was applied to conduct the HOTS approach and is supported by some principles of L2 learning. First, the most effective way to learn a language is through using it as part of interaction with and conveying meanings to others (Richards and Rodgers 2001; Williams 1998). Language learning is believed to be motivating when students are focusing on something other than language, such as making an evaluation or analysis; that is, when students use the language as a means of acquiring information or expressing their own ideas rather than as an end in itself, they learn the language more successfully.

Second, an activity is purposeful if it involves the learner in using language that not only conveys meaning, but also contains some value to the learner; this may be an educational value, or a value related to enjoyment, interest, need or sense of fulfilment, such as, expressing beliefs and ideas (Williams 1991). It is through using

thinking skills to evaluate or to reason that the value to the students themselves is realised. The function of conveying meaning, acquiring information, and putting learning to work while interacting with others might be carried out more thoroughly through incorporating elements of thinking skills into L2 learning material.

Spear and Sternberg (1987) argued that using thinking-based questioning is considered to be the most suitable means for the teaching of thinking. It encourages oral interaction between the teacher and students and creates the space to think. Therefore, this study will integrate thinking skills into L2 learning through thinking-based questions; students can learn the target language and develop thinking spontaneously.

To formulate questions, Bloom's taxonomy of cognitive domain (1956) is often applied as a thinking framework due to its systematic categorization of thinking levels which is easy to adopt. It is a useful model for developing material to help students to learn to think and it has been widely adopted to formulate activities (Adams-Smith 1981; Waters 2006) and higher-order questions both in L1 and L2 classrooms (e.g. Alcon 1993). Therefore, this study adopted Bloom's taxonomy as the fundamental thinking framework to formulate questions so as to scaffold students to think critically. The following will further discuss this taxonomy in depth.

2.4 Bloom's taxonomy

Bloom's taxonomy of cognitive domain (Figure 5) is a system of classification in which six levels of complexity of human thought are identified, namely, from the least to the most complex: knowledge, comprehension, application, analysis, synthesis and evaluation. This taxonomy further classified these six thinking levels into lower-order and higher-order thinking, as discussed in the following sections.

	Cognitive levels	*Thinking skills*
Lower order thinking	Knowledge	Remembering, recalling
	Comprehension	Summarizing, explaining
	Application	Transferring, classifying
Higher order thinking	Analysis	Reasoning, inferring
	Synthesis	Originating, creating
	Evaluation	Judging, assessing

Figure 5. Bloom's taxonomy of six thinking levels

2.4.1 Lower-order thinking

Knowledge, comprehension and application are categorized as lower-order thinking. In L2 classrooms this is fundamental thinking used to acquire the linguistic competence like memorizing, understanding the text and imitating (Ding 2007). This type of thinking is often utilized for rote learning and answering lower-order or display questions which require students to recall facts, explain or summarize (Long and Sato 1983; Tan 2007). Long's interaction hypothesis (1983) also requires students to use this thinking when negotiating for meaning like confirming or clarifying the meaning. However, the use of lower-order thinking is what the Taiwanese HE aims to discourage. As a result, the government has embarked on higher-order thinking promotion in learning.

Lower-order thinking skills are crucial pre-requisites for the higher skills and they form the foundation of all other skills. With the HOTS approach, lower-order thinking is essential for students to activate their schemata so that higher-order thinking can be exercised. Lower-order thinking skills are not disparaged, but educators should not be content with their students gaining mastery over the lower-order thinking skills alone; if

the instruction addresses only these levels, there is no basis for a student to learn the knowledge proactively and creatively and to transfer the skills to novel situations.

2.4.2 Higher-order thinking

Analysis, synthesis and evaluation are classified as higher-order thinking which can be seen as high cognitive thinking. This type of thinking is mainly used for active learning where students need to think critically and creatively. The use of higher-order thinking in learning is promoted by Taiwanese HE.

Higher-order thinking is mainly operated when answering higher-order questions. However, in the process of using higher-order thinking, lower-order thinking is also operated as a pre-requisite of higher-order thinking. As Bloom (1956) suggests that value or judgment (evaluation thinking level) cannot be applied until one knows the facts, understands the facts, can apply the facts, is able to take the facts apart and put the facts back together.

It needs to be pointed out that the categorization of thinking level is also context dependent (Barden 1995). For example, if the ideas elaborated like giving reasons or evaluating (higher-order thinking) can be found in the text, is taught by the teacher or is previously articulated by the other student, then it becomes a lower-order thinking idea because the speaker is simply recalling the memory. In contrast, if the ideas elaborated, for example, listing 10 animals that are mammals (example extracted from Barden 1995) which appears to be like lower-order thinking, cannot be found in the textbook, is not taught by the teacher, or is not previously articulated by other students and it requires higher-order processing, then this idea is categorized as higher-order thinking. In addition, if a student critically evaluates or comments on certain social events under the circumstance that he/she has already discussed this issue earlier in another context, for example, outside the classroom, this type of idea is deemed to be

lower-order thinking. However, it is sometimes difficult to trace back whether a student answers a particular question by using a high-level cognitive process or the relatively low-level process of knowledge recall, as acknowledged by Bloom (1956). In the present study, cognitive levels used by students will be assessed and categorised based on the discussion context. Reasons to reach this decision are first, the variable of different interlocutors that promotes a very unique context which will not occur elsewhere, and second, it is hard to identify whether students have encountered the topics before. In addition, when a higher-order thinking idea produces for the first time, this is considered as a higher-order thinking idea, whilst this idea is repeated for a second time, then it becomes a lower-order thinking ideas.

The main strength of Bloom's taxonomy is that it is logical and hierarchical, guiding the educator in a process leading from the most simple to the most complex form of cognitive skills. It is also a great advantage that Bloom links the mastery of these skills to particular behaviours and supplies numerous evaluative techniques (Bloom 1956). However, some scholars, such as Paul (1993), criticize the taxonomy because of its misuse as a 'cookbook.' Although claims of misuse do not logically implicate the taxonomy itself, it is an important warning to those who would apply it. As argued by Morgan and Saxton (1991, 1994), questions, which generate thinking, are supposed to follow the same hierarchical structure from the simple (recall) to the complex (evaluation) as a structure, yet the taxonomy is not a constructive way of planning and asking questions. In the present study, students could be asked certain lower-order questions in order to activate their background knowledge and then higher-order questions were posed. The questions asked were not in any order. Rather it was possible that the teacher applied any type of questions based on the context required.

2.4.3 Difficulty and complexity

Bloom's taxonomy distinguishes the difference between difficulty and complexity. To look into the taxonomy in more detail, complexity describes the thought process that the brain uses to deal with information, namely, the six thinking levels mentioned above. Difficulty refers to the amount of effort that the learner must expend within a level of complexity to accomplish a learning objective. To put these two abstract concepts into concrete examples, the complexity level increases when students move from lower-order thinking to higher-order thinking, for example, from comprehension to evaluation. The difficulty level increases as students manipulate a greater amount of information within a certain thinking level, for example, students manipulate a longer article while using analysis thinking. This implies that both the thinking levels and the amount of effort required need to be taken into account while developing students' cognition.

If teachers are to raise learner thinking, they need to increase the complexity rather than difficulty to guide learners into a more complex thinking level (Sousa 2001). However, it is very important to note that, as Sousa (ibid) argues, along with Bloom, the real connection to ability is *difficulty*, not *complexity*. In the learning process, fast learners can learn a concept quickly and, in the mean time, their brains group the concept's sublearnings into important and unimportant categories. If the teacher attempts to move up the taxonomy, these fast learners have the more important attributes of the concept in their working memory to use appropriately and successfully at the higher levels of complexity. Slower learners, if they do not sort the information properly when receiving a new concept, might clutter their working memory with all the sublearnings (important and unimportant), and may not be able to recognize the parts needed for more complex processing. Teachers who recognize the difference between difficulty and complexity can help slower learners improve their thinking and achievement significantly (Sousa 2001).

The discussion of difficulty and complexity implies that students' ability in handling larger amounts of information with HOTS is an indicator of higher intelligence. It indicates that the difficulty level of HOTS can be determined by the frequency of HOTS: the more students produce HOTS, the higher their cognition is. However, the context where students use their high cognition can be a factor which influences the effort expended. For instance, two students produce the same frequency of HOTS: one is in a context of merely answering a higher-order question, and the other is in a context of social interaction. The effort expended for producing HOTS in the latter context is greater than the former because it requires students to process more information with HOTS when students are exchanging the ideas. In the present study, different thinking tasks require different level of HOTS: for instance a thinking task can mainly require students to use analysis and evaluation, thus it can be difficult to evaluate students' ability of creative thinking; the decision of not examining students' cognitive development by the complexity level is reached. Therefore, students' thinking development will be determined by the difficulty level: the frequency of HOTS and how HOTS are used in the context.

2.5 Higher-order questions

Having explored higher-order and lower-order thinking and application to L2 contexts and this study, it is now important to consider the type of questions that scaffold (Wood et al. 1976) the move for students to higher-order thinking. This section moves on to explain what higher-order questions are and how to formulate such questions. It then looks at the debate about the effectiveness of these questions in L1 classrooms. The literature of higher-order questions in L2 classrooms is rather limited and so there is no existing debate on this issue. Instead, the potential effectiveness of higher-order questions in L2 learning is explored.

2.5.1 Formulating higher-order questions

Questions can be ranked in order according to the level of thought required for the response (Bloom 1956; Costa 2001; Wilen 1987; Marzano et al. 1988). The most common hierarchy in ranking the cognitive level of questions is Bloom's taxonomy (Section 2.4) with six categories of cognitive responses. With Bloom's taxonomy, there are also various categories of higher-order questions used among researchers. For example, Winne (1979) defines higher-order questions as those which require learners to manipulate information with application, analysis, synthesis and evaluation, while lower-order questions are those which demand only knowledge and comprehension. Higher-order questions can be referred to as high cognitive or divergent questions. Lower-order questions are sometimes referred to as low cognitive or convergent questions. In the current study, higher-order questions are defined as those which require students to manipulate information with analysis, synthesis and evaluation (as higher-order thinking in Bloom's taxonomy), while lower-order questions are those which demand knowledge, comprehension and application (as lower-order thinking in Bloom's taxonomy).

Based on Bloom's framework, Morgan and Saxton (1991, 1994) and Sanders (1966) formulated questions for classroom use and described the thinking skills applied in each thinking level. Higher-order questions formulated in this study were mainly based on these question stems. The following demonstrates phrases and examples of questions for each thinking level (adapted from Morgan and Saxton 1991, 1994).

Knowledge:

Who? What? When? Where? List....

Where can we find killer whales?

Who won the game?

Comprehension:

What is meant by...? Can you rephrase...? Can you describe...? Explain...

What is the main idea of this article?

Can you describe what happened to him yesterday?

Application:

What would happen if..? If you were...? What is a new example of...? How is...related to..? Do you know another instance where...?

If you were him, what would you do?

Do you know another instance where people dine with hands?

Analysis:

Why? What was the purpose...? What conclusions can you draw about...? What are the strengths and weaknesses of...? What is the difference between... and...? How is... similar to...?

Why is it important to make students support their inferences and conclusions?

Can you compare your project with that presented yesterday?

Synthesis:

How could we...? How can...? What if...? I wonder how...? What would happen if...? What do you think will happen if..? What is a possible solution to...?

What is a possible solution to this problem?

How can we create our own itinerary based on the budget and safety?

Evaluation:

Which is better? Would you agree that...? Would it be better if...? What is your opinion...? How would you feel if...? Is it a better solution to...? Can you defend your position about...?

Is it a suitable travelling package for senior citizens? Why?

What is your opinion towards backpack travelling?

Having stated the formulation of higher-order questions, it is important to know further criterion used to determine the level of questions applied in classroom context. Barden (1995) argued that higher-order and lower-order questions are context dependent and influenced by the objectives or expected learning outcomes of the lesson. That is, if the answer to the question is not taught by the teacher and is figured out by students themselves, then this question is defined as a higher-order question, even though the question appears to be lower-order in Morgan and Saxton's (1991, 1994) formulation.

Conversely, if the answer to the question is taught by the teacher and students just recall their memory to answer the question, then this question is identified as lower-order, even it is a higher-order question in nature.

2.5.2 The debate about the effectiveness of higher-order questions

Several researchers have claimed that higher-order questions conducted orally by teachers have positive effects on student achievement in the L1 classroom. Empirical studies have revealed two findings: first, there is a strong positive relationship between the cognitive level of questions and the cognitive level, length and syntactic complexity of student responses (Willson 1973; Cole and Williams 1973; Arnold et al. 1974; Smith 1978; Gall et al. 1978; Yang et al. 2005); second, the use of high cognitive questions helps students gain a better understanding of subject content (Ryan 1973). In addition, Redfield and Rousseau (1981) while critically reviewing the effect of studies conducted in the 1969s and 1979s, confirmed that regardless of the type of study or degree of experimental validity, teachers' predominant use of higher-order questions did have a positive effect on student achievement.

There are, however, a number of researchers (Gall 1970; Rosenshine 1971, 1976; Dunkin and Biddle 1974; Gall et al. 1978; Winne 1979; Mills et al. 1980; Dillon 1981; Samson et al. 1987) who claim that the effect of using higher-order questions on promoting student achievement, thinking and discussion has persistently been noted as not as effective as claimed. In particular, Winne (1979) pointed out that some research conducted during the 1960s and 1979s has significant methodological imperfections, for instance, a lack of delayed post-test to show the retention, a lack of replicability due to insufficient information, insufficient description of dependent variables. Indeed, the research design can affect the validity and reliability of the study. For a credibility of the present study, issues like replicability and retention will be considered and independent variable will be controlled and clearly described.

2.5.3 Higher-order questions in L2 classrooms

Question generation involves concentrating on main ideas while checking to see if the content is understood (Palinscar and Brown 1984). Scardamalia and Bereiter (1985) suggest that question generation is one component of teaching students to use higher-level cognitive functions independently. The change from behaviourist to cognitive theories of learning has highlighted what the learner does and how the learner processes information during the lesson rather than focusing on what the teacher does. From the late 1980s, the research conducted in the field of higher-order questions in L1 classrooms had shifted from teacher questioning to training students to ask high cognitive questions and to examine the training's effectiveness (King 1989, 1990; Foote 1998). This trend has affected research in the field of higher-order questions in L2 classrooms. Alcon (1993) was the first to study the process of training students to generate higher-order questions in a foreign language and found that it promoted the kind of verbal interaction which facilitated students' understanding and written production of the foreign language. It needs to be mentioned that the literature of higher-order questions reviewed in the following section includes both teacher questioning and students questioning. Regardless of who posed the questions, this study focuses only on the effect of higher-order questions.

The effect of higher-order questions on students' speaking and thinking and students' attitudes towards higher-order questions will be examined in the next section. The research gap will then be critically identified. This helps to inform the formation of research questions in this study. The critique of questioning behaviour will inform the design of a conceptual framework. One issue that needs to be considered before looking at the effect of empirical studies is whether students' utterance can demonstrate their thinking conducted in mind. According to Givens (1976 cited Mantero 2002, p. 441), discourse is one way that intellectual functioning is exhibited in

college-level classrooms. It is impossible to examine students' thinking process, but it is possible to observe students' thinking through verbal behaviour.

Empirical studies (Brock 1986; Nunan 1990; Alcon 1993; Godfrey 2001) show that the length of students' response to referential and higher-order questions is greater than that to display and lower-order questions. This finding indicates that opportunities for talk increase when learners are required to express their opinions.

The syntactic complexity of students' utterances has also been shown to be significantly higher when answering referential and higher-order questions than display and lower-order questions (Brock 1986; Nunan 1987, 1990; Godfrey 2001). It shows that when learners are required to express their ideas, more complex syntax can be triggered and thus develop linguistic competence.

The above studies show that with higher-order questions students can talk more with greater grammatical complexity. Nevertheless, it is necessary to give a thought for the type of talk occurring in such questioning behaviour. As Rowe (1974) stresses, high frequency of questioning behaviour makes the talk feel like an 'inquisition' rather than a reasonable conversation. Besides, such question-only teaching or learning might not be practical for a classroom where students have different learning styles. Dillon (1979) argues that, a variety of teaching techniques may readily be substituted for questioning with no loss of effect on achievement. This informs the design of the conceptual framework of the present study where teacher questioning will not be used as a main technique to carry out the innovation. In addition, student talk should not be like an 'inquisition', rather it needs to be reasonable conversations, as the types of talk occur in a social context (as discussed in Section 2.2.1).

An exploration of the length of utterance and grammatical complexity informed only students' linguistic competence (knowledge of language); the extent to which students can achieve a good command of spoken language (use of language) is not yet explored. The ability to communicate in a social context is very important for language development as stated in Vygotsky's social constructivism and it is also essential for second language acquisition as argued in Long's interaction hypothesis. Therefore, this study will further investigate students' speaking performance in a social context.

Research (Ayaduray and Jacobs 1997; Godfrey 2001; Ali and Daud 2003) has shown that students' responses demonstrate a much higher cognitive level in responding to higher-order questions than lower-order questions, where the frequency of high cognitive thinking increased: students elaborate the reasons, describe how to solve the problem, and use explicit criteria to evaluate. Display and lower-order questions mainly elicit replies from memory, which are not high cognitive demanding, whilst referential and higher-order questions tend towards higher cognitive demands, where learners are required to use reasoning, analyzing, evaluating, justifying, and creative abilities. The findings of empirical studies reveal that higher-order questions require high cognitive thinking and such practice enhances students' cognition.

The empirical studies reveal an increase in the frequency of high cognitive thinking through higher-order questioning. However, it needs to be questioned whether such questioning behaviour genuinely facilitates thinking. As Dillon (1978) and Yamada (1913) argue teacher questioning makes students passive by depriving them of opportunities to think independently and critically, to explore further, to solve problems and to enquire; such behaviour constrains student contributions to the learning process (Fairclough 1989). This discussion implies question-only teaching is not an appropriate means to allow students to think proactively. As previously discussed,

Dillon (1979) suggests a variety of teaching technique can be applied, for instance, group discussion.

In addition, cognitive improvement explored in these studies was determined by the frequency of high cognitive thinking only; an exploration of how high cognitive thinking is used in a social context will lead to a better understanding of cognitive development, as discussed in Section 2.4.3. Vygotsky's social constructivism and Swain's output hypothesis have pointed out the importance of social interaction in facilitating cognitive and L2 speaking performance. Therefore, this study will further investigate how HOTS are used in a social context: whether they are used for merely answering the question posed or commenting on others' thoughts.

It is worth noting that teacher questioning is not always able to elicit responses successfully from students (Wu 1993). Students might not produce answers, regardless of the type of questions asked, due to insufficient wait-time, students' own attitude towards questioning and a lack of risk-taking disposition. When the wait-time is not sufficient, L2 students rarely engage in conversations or they just fail to provide a response to the question (White and Lightbow 1984; Wu 1993; Tan 2007). A lack of risk-taking disposition, as found in the exploratory study, and negative attitudes (Wu 1993) can also inhibit learners from using the target language. Such kind of non-response or restricted response can be facilitated by giving sufficient wait-time or, as Wu (1993) found, by probing. Arguably, these two strategies can facilitate the use of higher-order questions. The studies of higher-order questions reviewed above did not include these two strategies into their research design. Therefore, the present study will adopt wait-time and probing as parts of the innovation to facilitate learning. These two strategies will be further explained in Section 2.6.

Attitudes towards higher-order questions have been shown to be both positive and negative. Tan's (2007) study of teachers' questioning behaviour found that higher-order questioning behaviour might have a negative affective impact on university students in China. Students complained that they were not able to manage higher-order questions because they felt that it was hard for them to think in depth in front of the class and felt rather embarrassed. Steven (1912 cited Dillon 1981; Santiesteban 1976) argued that high rates of teacher questioning might make students suffer from nervous tension. In contrast, Lin (2005) found that undergraduates in Taiwan are motivated in learning English and critical thinking through films by being asked higher-order questions. This finding indicates a high possibility that Taiwanese students might have positive attitudes towards the HOTS approach.

Carlsen (1991) claimed that in order to determine the valid level of questions, not only the effectiveness of the questions should be considered, but also the related influence of questioning behaviour on students' performance and students' attitudes towards it. Thus, students' attitudes towards and perceptions of this innovation will also be explored.

2.6 Strategies relevant to questioning

Wait-time and probing are strategies which facilitate questioning. In a social constructivist approach teachers can use probing to scaffold or mediate students' thinking. To allow students' thinking to be processed successfully, wait-time is essential. In this section, studies of wait-time in L2 classrooms and its relationship with higher-order questions are presented. Probing then is examined.

2.6.1 Wait-time

Wait-time is often linked with questioning bahaviour. Most studies distinguish wait-time between post-question and post-response (e.g. Kissock and Iyortsuun 1982; Duell et

al. 1992). According to Rowe's (1980) operational definition, post-question wait-time is the time between a teachers' question and a student's response, while post-response wait-time is the time between a student's response and another student's response or the teacher's resumption of speech. Post-question wait-time allows students to form an answer and to respond while post-response wait-time allows other students to reflect on the ideas contributed or to elaborate their own opinions. Higher-order thinking is more cognitively challenging than lower-order thinking in nature, and therefore, a demand of wait-time for processing information with higher-order thinking increases.

Research has shown that L2 students are rarely given enough time - an average post-question wait-time is of two seconds - to think and formulate their answers before teachers rephrase their questions or ask another student to answer (White and Lightbown 1984; Godfrey 2001; Tan 2007). It was found that when teachers wait from three to five seconds after asking a question, there is more participation by more students (Nunan 1990) and more accurate responses (Meredith 1978).

Based on research in language acquisition and on experience of teaching in a program which emphasizing communicative interaction, White and Lightbown (1984) recommended that the ideal post-question wait-time is five to ten seconds. Regarding the post-response wait-time, Godfrey (2001) found that in university L2 classrooms the post-response wait-time was of 0.80 seconds, while Shrum (1985) found that in high school L2 classrooms the post-response wait-time was 0.73 seconds. These two figures are far less than the recommended post-response wait-time in L1 classrooms (Rowe 1986), which is three seconds. In a teacher-led setting sufficient post-response wait-time indicates a greater opportunity to allow students to interact with each other and dominate the talk. This is a very important element which needs to be incorporated into this study due to one aim of this study being to encourage students

to interact with each other in a teacher-led setting. Therefore, not only the post-question but also the post-response wait-time will be applied into the design of this innovation.

Many studies of L1 and L2 classroom have documented that a longer wait-time is associated with higher-order questions (Arnold et al. 1974; Tobin 1987; Duell et al. 1992; Stahl 1994; Godfrey 2001). Tobin (1987) claims that the extended wait-time appears to facilitate higher cognitive level learning by providing teachers and students with additional time to think. Acquisition can be facilitated when wait-time is increased. Therefore, to encourage higher cognitive responses, adequate wait-time is essential (White and Lightbown 1984; Tsui 1996). Godfrey (2001) investigated teachers' higher-order questioning behaviour in both L1 and L2 classrooms at a university in America, and suggested L2 students need more time than L1 students in order to produce the longer and more complex responses that higher-order questions elicit. When wait-time was insufficient, L2 students were likely to fail to respond.

Having considered the importance of wait-time in cognitive development and student interaction, this study will apply a minimum five seconds for post-question and post-response wait-time. It needs to be pointed out that the exact wait-time needed also depends on the practical situation when conducting this innovation. Therefore, wait-time conducted can be longer or shorter than what it is designed to be. In addition to wait-time, probing is another strategy which facilitates the elicitation of more utterances from students and allows further opportunities for thinking. The following section will examine what probing is and its functions.

2.6.2 Probing

Probing is a questioning strategy, which facilitates the eliciting of students' responses. Probing means that a question is followed up by one or more supplementary questions

so that the teacher can elicit additional responses from a student. Higher-order questions alone may not guarantee responses with explanations or logically reasoned evidence. One function of probing is to allow the teacher to search for the reasoning based on the response made by the student. In Wu's (1993) study of L2 classroom interaction and teacher questions, students' responses were usually restricted irrespective of the types of question that elicited them. However, it was found that when the teacher probed for their reasons, students produced longer and syntactically more complex answers. Blosser (1973) claimed that the use of probing could facilitate discussion and generate more verbal behaviour.

Another function of probing is to help teachers to expand students' ideas. Some ideas given by students may be interesting and there is value in exploring them further. When students give different ideas and a teacher does not facilitate expansion of students' ideas, or search for their reasoning, even though the question may be open or higher-order, a shared and co-constructed sense of the meaning is lost (Smith and Higgins 2006). In contrast, if a teacher can expand students' ideas with the probing strategy, students will have more opportunities to elaborate their ideas and beliefs which are valued. Gall (1970) suggested that follow-up questioning of the student's initial response has substantial impact on student learning. In the present study, probing will be used in a teacher-led setting to model to students how to probe for further information in response to the information given. The questions used for probing were mainly higher-order questions in this study, but occasionally lower-order questions could be applied.

2.7 Theoretical conceptual framework

The literature reviewed informs the theoretical conceptual framework of the Higher-Order Thinking Skills (HOTS) approach designed for the current study. The rationale and an explanation of the framework are now provided.

2.7.1 The rationale

Approaches which encourage thinking advocate scaffolding of students' thinking rather than a directive role for the teacher. The literature (Section 2.3.4) shows that the most suitable means for the teaching of thinking is through thinking-based questioning. However, as discussed in Section 2.5.3, a high frequency of teacher questioning carries a number of disadvantages for learning. Therefore, teacher questioning was used in a teacher-led setting as a part of the innovation to model the tasks. This framework contained two elements: teacher modelling and collaborative group work. The modelling is an essential part of the implementation of the thinking tasks. As Burns and Joyce (1997) assert, many communicative activities have focused on getting students to speak, rather than on providing them with the means to interact. It can be argued that without the means to interact, the thinking tasks may not be effective in L2 classrooms. It can be seen as the means of gaining skills required in group discussion.

The rationale for including collaborative group work is that firstly it is a way to decentralize classroom communication in an attempt to encourage more students to participate in interaction due to the difficulties of managing the turn-taking of a large number of students. Secondly, most students have high anxiety of speaking English in Taiwan (Chen 2002). It is necessary to provide a less stressful environment to lessen the anxiety for students to use the target language and conduct higher-order thinking and this can be done by promoting group work (Tsui 1996; Long and Porter 1985). Group work allowed students to model what they have learned and practice in groups. Without modelling or the thinking tasks, the HOTS approach would be liable not to work.

The rationale for adopting wait-time and probing was articulated in Section 2.6. Based on the studies reviewed, the present study adopted a minimum five seconds for post-

question and post-response wait-time. It needs to be noted that both wait-time and probing were mainly used in the modelling stage; students could manage their own wait-time and probing strategy in group discussion.

2.7.2 An explanation of the framework

The framework of the HOTS approach, as shown in Figure 6, starts with the idea of infusing thinking skills through questions; thus higher-order questions are used. At the modelling stage, the teacher first poses a higher-order question, followed by a minimum five-second post-question wait-time in order to allow learners to understand the question, get an idea and form the answer. Two possible types of response might occur after the wait-time: a restricted response or an elaborated response. A restricted response means a response which answers the question but with a very short utterance and without sufficient explanation or logically reasoned evidence. In contrast, an elaborated response refers to a response consisting of an explanation or logically reasoned evidence and involves a longer utterance with higher cognition. An elaborated response corresponds to analysis, synthesis, and evaluation in Bloom's (1956) taxonomy.

After posing a higher-order question, an elaborated response might occur directly, with the student providing a response with explanation and logically reasoned evidence. On the other hand, if a restricted response occurs, the teacher is required to use probing, asking a further question to search for possible reasons, in order to give the learner a further chance to talk and to guide him or her to elaborate more specifically. After probing, wait-time is also required. With probing and wait-time, the learner might be able to produce an elaborated response or might fail to produce one. However, the literature (Section 2.6.2) shows that there is a greater possibility that learners will be able to produce a more sophisticated response when the probing strategy is used.

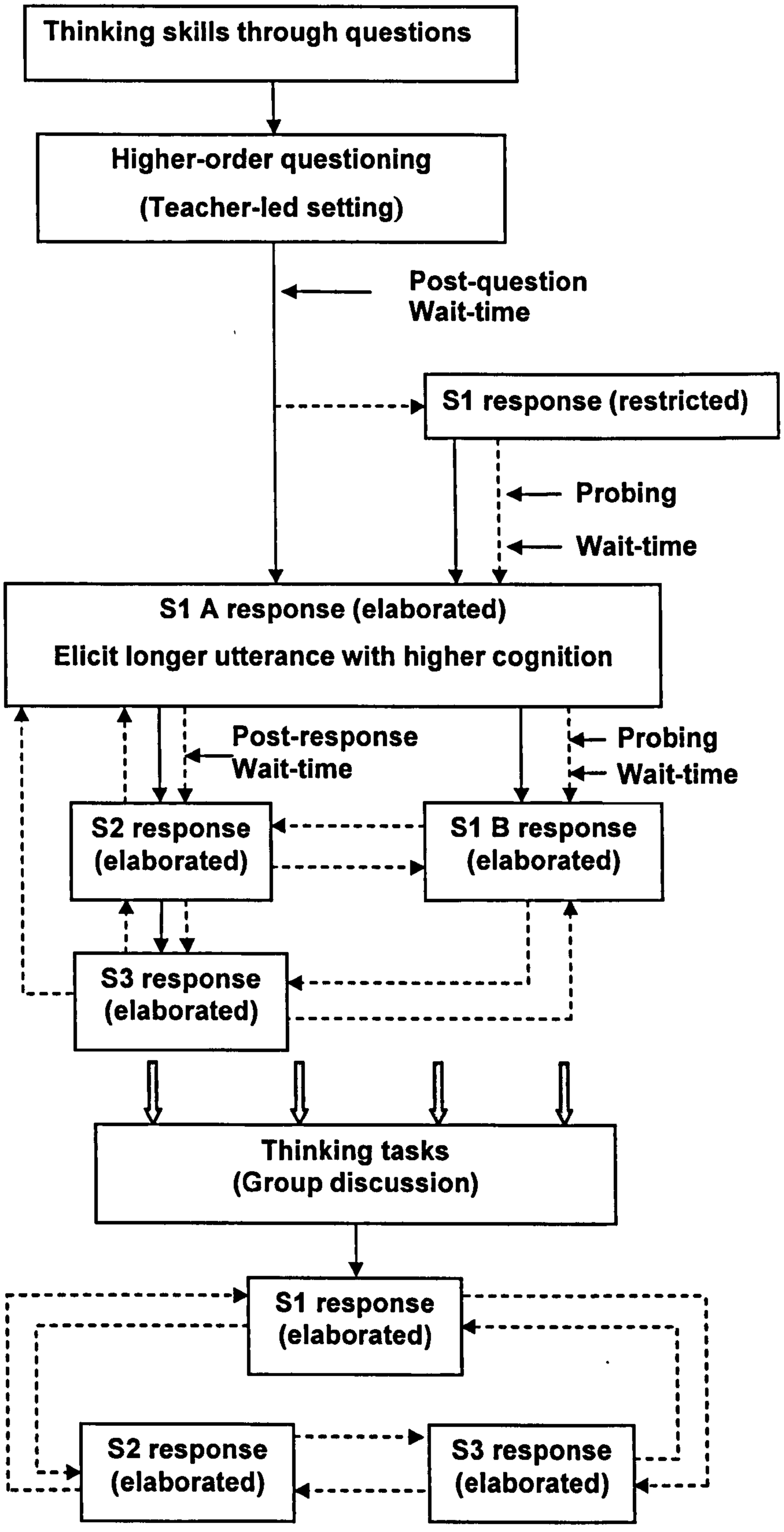


Figure 6. Theoretical conceptual framework

After an elaborated response (Figure 6, S1 A response) is elicited, it can be followed by two possible actions by the teacher. One action is that the teacher uses probing along with wait-time to expand upon the student's ideas, and then the student can elaborate further on the information or ideas (Figure 6, S1 B response). The other action is that the teacher gives five-second post-response wait-time in order to allow another student (Figure 6, S2 response) to reflect on the previous response or elaborate about his/her opinions and ideas. Followed by this, a further student (Figure 6, S3 response) can also reflect on the previous responses or elaborate about his/her opinions and ideas. The procedure from 'S1 A response' to 'S3 response' creates the 'interaction space' among students and is designed to allow students to talk freely and challenge each other's ideas. Within this interaction space, it is possible that socio-cognitive conflicts (Section 2.2.1) occur and this may help students to think critically and reconstruct their knowledge. Very importantly, teachers are not encouraged to use numerous higher-order questions within a given period of time, e.g. in one lesson, since the higher the number of higher-order questions, the lower the interaction among students can be.

Students then can model the above interaction pattern in group discussions through the thinking tasks. Students can probe for further information needed and comment on others' thoughts. Consequently, the discussion can be interactive, which may lead to more elaborated utterances and higher-order thinking skills. The teacher at this stage plays as a facilitator and scaffolds the discussion in groups. How this framework fitted into the research design will be further explained in Chapter 3.

CHAPTER 3 - METHODOLOGY

3.1 Introduction

Having identified the research gap and generated a conceptual framework, this chapter outlines the research design used to answer research questions. Before introducing the research design, it is necessary to outline the aims, research questions and the methodology. This will provide the foundation to describe the methods used in the field of L2 higher-order questioning so as to inform the research design.

3.2 Aims and research questions

The main aim of the current study was to evaluate the practicability of the HOTS approach in a Taiwanese university L2 classroom. The research attempted to answer one main research question with three sub-questions, based on the innovation conducted. The research questions were derived from the L2 speaking problems observed, the literature related to higher-order questions and the identified research gap to form the focus of the study. They are as follows:

The main question:

Is it practicable to use the HOTS approach to develop speaking ability and foster higher-order thinking skills in a Taiwanese university L2 classroom?

The sub-questions:

1) How does the introduction of the HOTS approach impact on students' oral performance?
2) How does the introduction of the HOTS approach impact on students' observed cognitive behaviours?
3) What are students' attitudes towards and perceptions of the impact of the HOTS approach?

3.3 A critique of methods used in the field

Methods used in the current study are directly linked to research within the field of higher-order questions and L2 speaking. It is necessary to reflect on empirical studies which have investigated this area and critically examine the methods applied in order to inform the research design.

With regard to conducting higher-order questions in L2 classrooms, no case study was done, but three pieces of experimental research in which the study design was more similar to this research were found: Alcon (1993), Ayaduray and Jacobs (1997) and Ali and Daud (2003). One weakness of these studies is that qualitative information was not taken into account. The focus of these studies was mainly on students' production, e.g. higher-order thinking and oral performance. Students' verbal interaction was audio-recorded. Data were coded and analyzed quantitatively. The outcomes were dichotomous, either significant or not significant. No qualitative information was delivered to enable the readers to understand the deeper aspects of the experiment, e.g. students' attitude, learning contexts, etc. This leads to weaknesses of quantitative research: the researcher may miss out on phenomena occurring and knowledge produced may be too abstract and general for direct application to specific local contexts and individuals (Johnson and Onwuegbuzie 2004).

A second language classroom is like a society where there are diverse learning styles and viewpoints regarding learning (Oxford 1990); therefore, outlining the effects through statistical analysis is insufficient to provide a complete and holistic report. In addition, Carlsen (1991) argues that to determine the valid level of higher-order questions, students' reactions and the context need to be taken into account along with student responses to questions. Brown (1989) also claims that, when evaluating a language programme, such as conducting an innovation in a second language classroom, all relevant information, including qualitative and quantitative data, is

needed in order to promote the improvement of the programme conducted and assess its effectiveness and efficiency, as well as the participants' attitudes within the context involved. Thus, a more comprehensive result can be provided for the readers.

A second serious weakness of these studies is the failure to address whether the effect of the treatment was long lasting. The three studies used pre- and post-tests or just a post-test; delayed post-test was not conducted in any of the studies. Pre-test and post-test design allows the researcher to determine the immediate effect of treatment but not to ascertain the extent to which the learning is retained. As Winne (1979) argued, the effect of higher-order questions on promoting student achievement, thinking and discussion was not as effective as claimed: there was no delayed post-test to show how students maintained the skills. In addition, Mackey and Gass (2005) state that the real question for studies of second language learning is the extent to which a treatment truly results in learning. Learning is a process that may begin with a particular treatment, but it is not always clear whether the effects are sustainable. Therefore, a delayed post-test is arguably necessary to obtain a wider snapshot of the innovation effects.

The validity of Ali and Daud's (2003) research, exploring the effect of higher-order questions on critical thinking, may be questioned due to the rigour inherent in the design. There are two reasons for this: first, it was a post-test only research design; there was no pre-test to determine whether participants in the two classes were of the same cognitive level; second, their study had two independent variables: one was the treatment 'higher-order questions' and the other was the 'discussion forum'. The experimental group received the training in higher-order questions and practiced in a small group discussion, while the control group participated in classroom discussion, a teacher-led setting. Speaking practice in different forums has different effects. As Tan (2007), Tsui (1996) and Wu (1993) argue, in a teacher-led setting students fear to

speak English and it can be difficult for them to think critically. In contrast, group discussion forum provides a less stressful environment for students to use L2 (Long and Porter 1985; Tsui 1996). It is difficult to say whether the effects on thinking were attributable to the treatment or the different discussion forums. Thus, the current study will not refer to Ali and Daud's (2003) research regarding the effects of higher-order questions on thinking in the discussion chapter, Chapter 5.

This critical review has informed the research questions of this study: to investigate both students' speaking and thinking performance including their attitudes towards using high cognition so that a more comprehensive understanding was presented. It has also informed this study to apply quantitative and qualitative research design: examine the effects of the innovation with quantitative research and investigating related in-depth information with qualitative research. The following section will turn to illustrate the methodological conceptual framework.

3.4 Methodological basis

The current research design was developed using understandings from the nature of qualitative and quantitative research, the weakness of previous research and the research question themselves. Figure 7 demonstrates the methodological framework. The preliminary exploration (Section 1.3) was completed to explore potential L2 speaking problems. This together with the review of educational goals in Taiwan (Section 1.4) determined the literature to be reviewed, and thus the aims for this research emerged. The theoretical conceptual framework (Section 2.7) and the research questions were derived from the literature of thinking and second language learning reviewed in Chapter 2. The following will go on to elaborate the rationale for taking a pragmatist's standpoint and adopting a mixed-method approach.

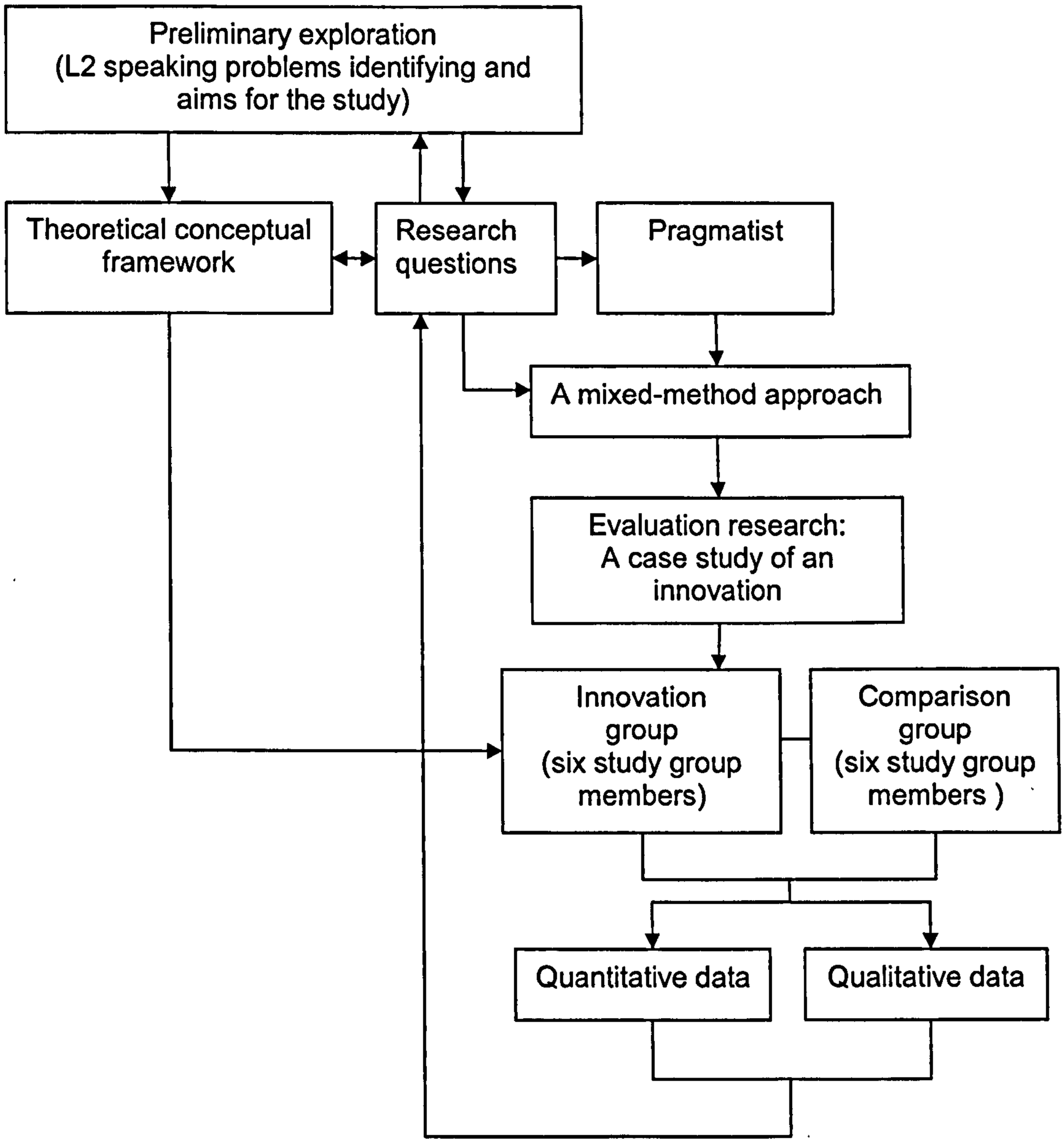


Figure 7. Methodological conceptual framework

One aim of this study was to seek solutions to solve L2 speaking problems. Pragmatism is problem-centered and real world practice oriented. Therefore, this study took the epistemological, ontological and axiological position of pragmatism. In order to understand the problem, all approaches, including quantitative and qualitative methods, should be used to explore potential solutions to the problem (Creswell 2003); pragmatists believe in the existence of both subjective and objective view points. They look to the 'what' and 'how' to research intended consequences. The truth is what

works for the problem at the time. In addition, the values held by the pragmatists play a role in the interpretation of the results (Onwuebuzie 2000).

Research methods should be guided by the research questions in that they should offer the best opportunities for gaining useful answers (Johnson and Onwuegbuzie 2004). The multifaceted nature of the research questions in this study and the review of methods used in the field of L2 higher-order questions studies meant the present study encompassed both quantitative and qualitative data, thus a mixed-method approach was adopted.

To appreciate a mixed-method approach, it is necessary to know both quantitative and qualitative research methods. Quantitative researchers usually try to generalize findings to a larger population and seek cause-and-effect relationships (Bryman 2001). Qualitative researchers, on the other hand, try to find in-depth explanations for particular events with a small sample (Muijs 2004). These two types of researchers have sometimes uncompromisingly different conceptual world views. They both assert the *Incompatibility Thesis* (Howe 1988), which posits that quantitative and qualitative research paradigms and methodologies cannot and must not be mixed because both paradigms operate under different ontological, epistemological, and axiological assumptions about the goal and nature of research (Tashakkori and Teddlie 1998).

Several researchers however argue that epistemological purity does not result in effective research (Miles and Huberman 1994; Onwuegbuzie 2000). Quantitative and qualitative studies each have their own weaknesses. One weakness of quantitative research is that variables relating to the phenomenon of interest might be omitted because of the focus on statistical generalization (Johnson and Onwuegbuzie 2004). In addition, the meaning of events to individuals is ignored, so that findings in quantitative research do not seem to connect with everyday contexts (Bryman 2004).

This means the knowledge produced may be too abstract and general for direct application to specific local contexts and individuals.

As to the weakness of qualitative research, a lack of objectivity is one major issue and the restricted scope of findings also causes problems of generalization (Bryman 2004). Both quantitative and qualitative studies have their own strengths, but due to their weaknesses, epistemological purity is arguably unable to provide a holistic understanding of the kind of phenomena studied in social sciences. As Onwuegbuzie (2000) stresses, the quantitative and qualitative division is counterproductive for advancing the field of the social and behavioural sciences. In addition, Creswell (2003) and Denzin (1978) assert that all data collection methods have limitations. For example, the bias inherent in any particular data source, investigators, and particular method can result in low reliability and validity. Therefore, quantitative or qualitative research alone is never able to overcome the limitations of each.

In order to gain a comprehensive understanding of educational and social phenomena, pragmatists (Onwuegbuzie and Leech 2005; Johnson and Onwuegbuzie 2004; Creswell 2003; Onwuegbuzie 2000; Newman and Benz 1998; Sechrest and Sidani 1995; Madey 1982; Sieber 1973) advocate epistemological ecumenism: integrating both quantitative and qualitative methods within a single study. Such integration is also known as a mixed-method or multi-method approach. Research methods used in a mixed-method approach are not as impeded by epistemological and ontological baggage (Bryman 2004) as pure quantitative or qualitative research.

Having explained the rationale for adopting pragmatism and the use of a mixed-method approach, it is now important to elaborate on the nature of this research which can be categorized as evaluation research and the rationale for applying a case study research design. Evaluation research involves the evaluation of innovations or

organizational programmes and usually attempts to determine whether the innovation has achieved its expected objectives (Bryman 2001). The typical approach in evaluation research is the experimental design. The current research design including an innovation and a comparison class was very similar to experimental design; however, it was a case study design. This study aimed to evaluate the applicability of the HOTS approach. It involves making judgements so that it can point out the efficiency, benefits or problems associated with this approach. To conduct evaluation research, Yin (2003) states that case studies can be adopted to explain, to explore or to be used as a meta-evaluation.

Yin (2003, p. 13) describes a case study as:

> 'an empirical inquiry that investigates a contemporary phenomenon within its real-life context.'

Case study researchers seek to understand a case holistically (De Vaus 2001; Gall et al. 1996). The main purposes for researchers in conducting case studies are to produce thick description, to develop possible explanations of the phenomenon and to evaluate it (Gall et al. 1996). Therefore, a case study method allowed the researcher to look into the HOTS approach in depth and retain the holistic understanding of 'how' it facilitated students' speaking and thinking performance in this particular context. One unique character of this type of design is that the result relies on multiple sources of evidence, with data needing to converge in a triangulating fashion, and as another result (Yin 2003). This matches the notion of a mix-method approach.

This study intended to provide *theoretical* generalization. As Yin (2003) argues, the heart of a case study is to test a theory regarding a particular phenomenon within its real-life context. The particular phenomenon in this study referred to the practicability of the HOTS approach. The context under studied was in a Taiwanese university L2

classroom where students' learning style was passive and thinking was not proactively operated, the class size was large, teachers dominated the instruction context and speaking sequence, and exposure to English use outside the language classroom was limited. Therefore, to examine the practicability of the thinking approach in Taiwanese HE context, a case study design was felt to be appropriate.

This study was designed as a case study of an innovation, and a comparison class was included in order to increase the validity; these two classes provided an overall picture of the effect of the HOTS approach. In each class six students were selected as study group members with video data collected to show students' speaking and thinking performance. The innovation class received the treatment drawn from the theoretical conceptual framework while the comparison class did not. Both quantitative and qualitative data were collected and analyzed based on the principles of a mixed-method approach. Results aimed to inform recommendation for enhancing L2 speaking and thinking and solutions for tackling L2 problems in Taiwan universities.

3.5 Research design

Having explained the methodology and rationale, this section will articulate how a mixed-method approach and a case study fitted into this research design to answer the research questions. In addition, an overview of data collection procedures will be illustrated.

3.5.1 A mixed-method approach

Hammersley (1996) classified three approaches to mixed-method research:

Triangulation: It refers to the use of quantitative research to corroborate qualitative research findings or vice versa.

Facilitation: This approach arises when one research strategy is employed in order to aid research using the other research strategy.

Complementarity: This approach occurs when the two research strategies are employed in order that different aspects of an investigation can be dovetailed.

Hamersley's classification provides guidelines as to how quantitative and qualitative research methods work together in order to answer research questions. This framework had an influence on this methodology: it helped the researcher to clarify the purpose of applying different research tools. For example, this study looked at students' speaking and thinking performance using the video data and these data also triangulated and complemented students' perceptions of their development in speaking and thinking observed in the interview and questionnaire data. The questionnaire data aimed to explore an overall view of students' attitudes towards the HOTS approach and they were complemented by the interview data where more in-depth information was provided.

According to Bryman (2007), a genuine integration of mixed methods aims to mutually illuminate the quantitative and qualitative components of researchers' analysis, interpretation, and writing up of their research. One critical issue relating to studies employing a mixed-method approach is that very few actually integrate the data during analysis (Bryman 2007; Green et al. 1989). Thus, the researcher in the present study had taken measures to ensure that the data were genuinely integrated at analysis, interpretation, and writing up stages (The framework for the analysis and interpretation of the data is presented in Section 3.8.).

The advantages of using a mixed-method approach were observed as follows. At the data analysis and interpretation stages the qualitative data helped to interpret, clarify and validate the statistical findings and therefore the effectiveness of the HOTS

approach. Conversely, the quantitative data facilitated the assessment of the generalizability of the qualitative data and shed new light on the qualitative findings. Thus, more complete knowledge was generated to inform theory and practice (Johnson and Onwuegbuzie 2004). This, on the other hand, enabled the researcher to better attend both to classroom processes and the contextual conditions that shape a learning environment (Johnson 1992). Students' learning process and contexts were illustrated clearly, for instance, the kind of difficulties like psychological, intellectual and speaking problems encountered, how they overcame or reduced these problems, the sort of learning contexts which facilitated their learning, in what way the thinking task that helped or did not help them learn and the impact of this study in their study and life. As a result, the researcher had greater faith in the findings with higher reliability and validity and therefore could make greater contribution to the field (Currall and Towler 2002).

Some disadvantages however to using a mixed-method approach occurred. It was challenging and time consuming for the researcher to learn about multiple methods and approaches, and understand how to combine them appropriately in order to reach the intended consequences (Johnson and Onwuegbuzie 2004). Prior to merging these two research methods it is necessary to understand the research per se, including its epistemological, ontological and axiological perspectives and data collection tools applied. It was a complicated process to triangulate and complement data at the data analysis stage and to interpret results at the interpretation stage. Analysing either quantitative or qualitative data individually is a rather straight forward process. For a mixed-method approach, it requires more than that: in order to use all the data collected to reveal the truth comprehensively, the researcher needed to find a way to weave both quantitative and qualitative findings together, rather than reporting them individually.

In addition, at interpretation and discussion stage it required the researcher to link the findings together and identify the relationship. For instance, the researcher needed to identify the emerging themes and how a particular theme impacted on other themes. The discussion had to bring the findings together into a cohesive whole. Such effort was far greater than simply discussing the individual findings found. This is what Bryman (2007) advocates that researchers need to devote more efforts for integrating the findings while applying a mixed-method approach.

3.5.2 A case study of an innovation

The case in a case study can be an individual, a group of people or an organization (Yin 2003). This research applied a comparative case study design: two cases with embedded units, as shown in Figure 8. One case was the innovation class which received the HOTS approach as an innovation, while the other case was the comparison class which received the teaching method similar to Grammar-Translation. This study aimed to investigate whether the HOTS approach was workable in the Taiwanese university L2 context by comparing with traditional teaching method, Grammar-Translation. Six study group members in each class were the embedded units where video data were collected. This case study lasted for 17 weeks (for detail, see Section 3.5.3, Table 2), focusing on the effect of the innovation conducted.

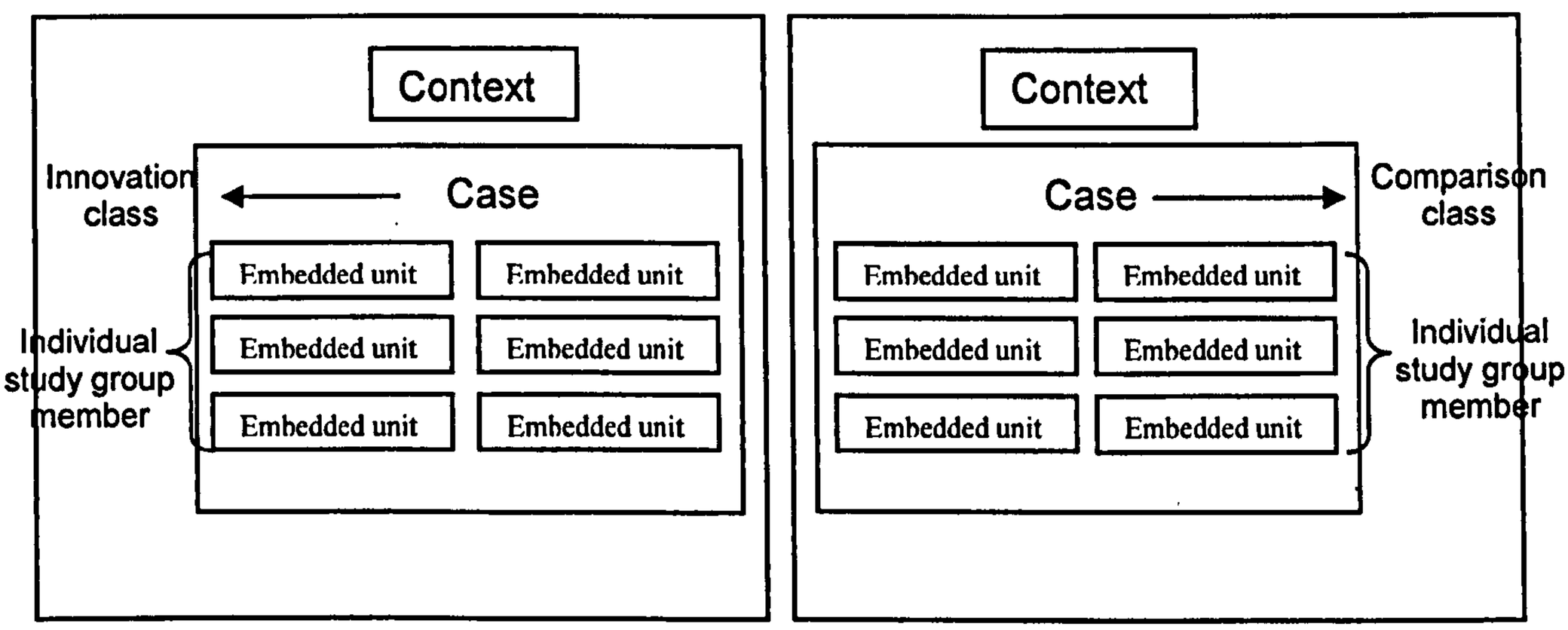


Figure 8. A comparative case study design

There were two main advantages of using a case study design in the present study. It allowed the researcher to focus on the individual (Mackey and Gas 2005) which referred to the six study group members. The researcher was then able to provide insights into the complexities of particular cases in their particular contexts. This, on the other hand, allowed the researcher to bring the study to life in a way that was not possible using quantitative research (Gall et. al. 1996). The thick description of the context helped to shed light on the complexities of the second language learning process.

Many methodological textbooks (e.g. Gall et al. 1996) state that the main disadvantage of case studies is the difficulty in generalizing the findings to other situations. Mackey and Gass (2005) also mention that case studies may provide valuable insights into certain aspects of second language learning, but single case studies are not easily generalizable. However, the current study aimed to provide *theoretical generalization* rather than *statistical generalization*.

The theory of higher-order questions has been proved effective in second language learning (Section 2.5.3) in some western countries and former British colonies. This study aimed to examine whether the theoretical conceptual framework (Section 2.7) worked in a Taiwanese HE L2 context. When the theory works in a different context with different participants, confidence with regard to the effectiveness of the theory increases.

3.5.3 An overview of data collection procedures

Table 2 illustrates the three phrases of data collection from the two classes: pre-test, post-test and delayed post-test. It also indicates the data collection tools employed in the different phases and from whom the data were collected. Pre-test was conducted in the first week, followed by a twelve-week innovation. Four weeks after the

innovation, the delayed post-test data were collected. The pre-test and post-test both collected data through questionnaires and video-recording from two classes; questionnaire data were collected from the whole class while video data from the study group members. The delayed post-test collected only video data from the study group members in the two classes. Six study group members and ten non-study group students in the innovation class were interviewed after the post-test. Data were collected mainly by the researcher in the period from March to June 2008.

Table 2. Data collection procedures

	Class	Data collected from	Pre-test Week 1	Post-test Week13	Delayed post-test Week 17
Video recording	I	Six study group members	√	√	√
	C	Six study group members	√	√	√
Questionnaires	I	The whole class	√	√	
	C	The whole class	√	√	
Interviews	I	Six study group members and ten non-study group members		√	
	C				

I= Innovation class, C= Comparison class

3.6 Participants and sampling

Before articulating the participants recruited in the preset study, it is necessary to point out that non-English majors were focused on in this study rather than English mjaors. There were two reasons for targeting non-English majors: one was that the majority of L2 learners in the Taiwanese university setting are non-English majors; the other was that they arguably have lower motivation to learn English and higher anxiety towards speaking English (Liao 2008).

The target learners were from a university in Taiwan. Seventy-two non-English major freshmen, aged between 18 and 19, from two classes were recruited for this study, using non-probability sampling. Selection of non-probability sampling was made on the researcher's judgement as to typicality (Robson 2002). The researcher built up a sample that was appropriate for the specific needs of the research: participants were targeted at university level with non-English major freshmen. The two classes were suggested by a lecturer in a university. These two classes had English lessons on the same day: the Maths class had a morning lesson, while the Visual-Art Design class had an afternoon slot. The original plan was that the formal lecturer would use her own teaching method to teach the comparison class in the morning and conduct the innovation in the afternoon. However, prior to commencing the innovation the lecturer decided not to conduct the innovation, instead the researcher took over this class (see full explanation in Section 1.5.3). Thus, the lecturer worked with the comparison class, while the researcher carried out the innovation with the innovation class.

The innovation class majored in Visual-Art Design while the comparison class majored in Maths meaning that two classes were of different majors. English language proficiency levels of the two classes varied from lower-intermediate to higher-intermediate. Approximately ninety percent of the participants in each class had passed the beginner level of General English Proficiency Test where the English proficiency is equivalent to being able to understand and use daily life vocabulary with approximately 2000 words and phrases stored. These two classes were both mixed-gender. It is worth noting that most students majoring in Visual-Art Design were female while the majority of students in the Maths class were male. Very few participants from both classes had been exposed to an English learning environment after school, for example, taking extra English lessons. The impact of having different subject majors in the comparative case study design was recognised and is discussed in Chapter 6. However, pragmatic reasons meant this was the most appropriate groups.

The study was completed as part of a compulsory general English course for non-English major freshmen. It was a two-credit course and offered two hours of English lessons per week. This course aimed to improve students' general English language proficiency and it lasted for one semester of 18 weeks.

Prior to commencing the study, ethical clearance was sought from the participants. At the start of the semester, the researcher explained the purpose of the research, what the research was about, and what participants needed to do in order to participate. A letter of consent and a consent form (Appendix 3) were handed out to both the comparison and innovation classes and all 72 students agreed to participate by signing the letter.

Within each class, study group members were selected on the basis of volunteer sampling (Milgram 1963). There were six study group members in each class where the video data were collected. The two groups' average scores in English language in the 1st semester were approximately the same. The average score in the comparison class was 83.5 while that of the innovation class was 84.5. Figure 9 shows the median of the score in two classes. None of these members took additional English lessons after school. A small sample was chosen because of the expected difficulty of managing a large amount of utterance data within the limited time. Before the study group members were chosen, the aim and the responsibilities of being a study group member, e.g. to be video-taped and interviewed, were explained. The formal lecturer of these classes suggested that, according to her knowledge about these participants, it was better for students to volunteer to join the study group and to form their own small groups of three; students' willingness to communicate has been shown to depend on the degree of acquaintance between communicators (MacIntyre et al. 1998). It was felt that group discussion might not work if the group members did not like or were not familiar with each other. Study group members in the innovation class

were chosen based on their volunteering, while those in the comparison class were chosen based on students' nomination. Although study group members in the comparison class did not volunteer to join the study group, their willingness to participate was confirmed through oral inquiry.

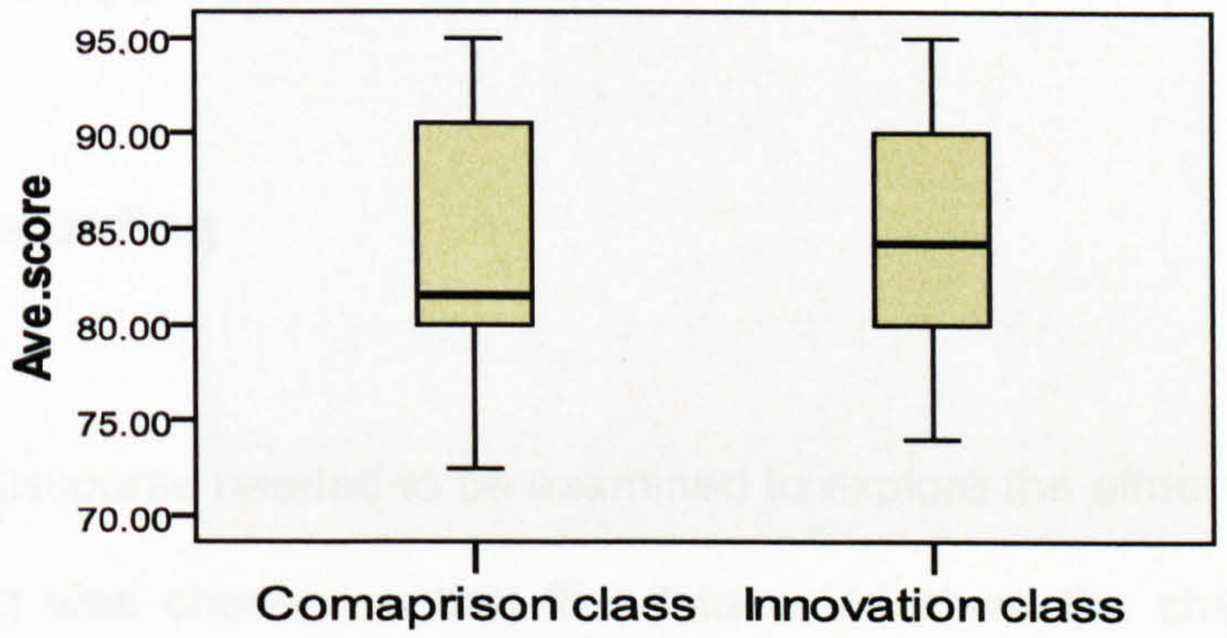


Figure 9. Median of study group members' average score

In addition, it is necessary to describe the role of the researcher and the relationship between the researcher and participants. The researcher, as an innovator, played the roles of both researcher and lecturer in this study: designing the tasks, organising things, telling the participants what to do, teaching, collecting the data, etc. In conducting this innovation, the researcher was involved in the participants' learning process. The relationship between the researcher and participants was an ongoing and evolving one in terms of teaching and learning. Sociologists often look at the sorts of relationship between the innovator or experimenter and subjects as power (Babbie 2004). Such power also exists in the hierarchical relationship between teacher and student in Taiwan; generally, teachers dominate the classroom while students do as they are told. This kind of relationship has the relevance for the potential outcomes of this study.

3.7 Justification of data collection tools

Having stated the sampling techniques and participants, this section justifies the research tools chosen, namely video recording, interview and questionnaires. The rationale for adopting each data collection tool, the design of the research tool, and the process of collecting the data are described in detail.

3.7.1 Video recording

Rationale

Students' oral discourse needed to be examined to explore the effect of the innovation. Video recording was chosen so that the data could show the change in students' utterances before and after the innovation. There are other means to examine students' utterance like using systematic observation. However, two study groups in each class were discussing simultaneously and video-taping would allow the researcher sufficient time to examine students' speaking. These data were mainly used for answering research sub-question 1 and 2 which aimed to explore students' speaking and thinking performance and also served as a function of triangulation with other data.

Design

Students' oral discourse was collected from the study groups during group discussion. A 15-minute discussion task was used in all three data collection phases: the 5Ws, but with different topics (for tasks used in Pre-, Post-, and delayed post-test, see Appendix 4). The number of questions was approximately the same in each task. Topics chosen were based on the related content in the textbook. For example, some units in the textbook were about personal characteristics, thus the topic 'What characters are the most important for you while choosing an ideal mate?' was developed.

Two camcorders were used to record the discussion of the two study groups. The camcorders were placed 1.5 meters away from the group of three. In addition, as the classroom noise level could affect the clarity of recording in such a big class size with cramped space, a voice recorder was placed on the table of each study group as a backup.

Process

Before conducting the group discussion the researcher explained what the task was about, what students needed to discuss. Participants were divided into groups of three, which were formed by the students themselves. When study group members were ready, the camcorders and voice-recorders were started. The researcher did not interfere in the group discussion, so the oral discourse was genuinely activated and produced by the study group members themselves. In the mean time, the researcher walked around the classroom and observed the discussion in other groups. When the time was up, video- and audio-recording stopped.

One issue occurred during the recording. A voice-recorder was placed on the table of each study group and after the researcher started the voice-recording, two groups stopped it on their own for the sake of planning for discussion at pre-test; it was late when the researcher noticed this. The voice quality in the camcorder was not good enough. Thus, the total recorded length gained from the voice recorder was shorter than 15 minutes.

3.7.2 Questionnaires

Rationale

The questionnaires attempted to explore students' general perceptions of thinking skills used and to investigate students' self assessment of their ability to adapt higher-order thinking in the L2 classroom before and after the innovation; these data mainly

answered the research sub-question 3 and were also used to triangulate and complement the video and interview data. The rationale for using questionnaires as a data collection tool was that they provide a relatively simple and straightforward approach to study attitudes, values, beliefs, and behaviour (Robson 2002; Dörnyei 2003) and they were time-efficient (Cohen and Manion 1994). However, it should be recognised that the data can be affected by the characteristics of the respondents (Robson 2002), e.g. their memory and experience. As with most methods there are advantages and disadvantages to using questionnaires for data collection; the design and administration process attempted to minimise potential impact.

Design

The questionnaire, in line with the innovation, was developed using theory based on the six thinking levels of Bloom's Taxonomy (Section 2.4). For ease of analysis and to offer more choices, the questionnaires used a 6-point Likert scale (Likert 1932). According to Dörnyei (2003), 5- and 6-point scales are the most common steps used in Likert scales with either an inclusion or exclusion of a middle category, e.g. 'neutral' or 'not sure'. An inclusion or exclusion of a middle category does not affect the relative proportions of the expressed opinions and does not modify the results significantly (Robson 2002). Thus, in order to impel students to specify their opinions a 6-point Likert scale without a middle category was adopted in the study, in which 1=strongly agree while 6=strongly disagree.

There were three sections in the pre-test questionnaire (Appendix 5). Section 1 aimed to explore students' attitudes towards the importance of thinking skills used in English classes, namely knowledge, comprehension, application, analysis, synthesis, and evaluation, which answered research sub-question 3. Three questions were designed for each thinking level, with a total of 18 items. Some of the concepts associated with thinking levels might be rather abstract for the respondents to understand; this is one

of the limitations of using questionnaires as a research tool. To make the questions more concrete for the respondents, an additional example was included. Section 2 aimed to explore how students assessed their own ability in using the six thinking skills to answer questions in English class. These data complemented and triangulated the video and interview data. The principle of setting the questions in Section 2 was the same as for Section 1. All18 questions in Section 2 contained an additional example. Section 3 asked about respondents' attitudes towards questioning-answering behaviour in the English class, which answer research sub-question 3 and were also used to complement and triangulate the video and interview data. Nine questions in this section were designed based on the literature of classroom questioning-answering behaviour (see Chapter 2), e.g. wait-time, probing and preference of intellectually challenging questions, etc.

In addition to the three sections mentioned above, one more section, Section 4, which asked about participants' opinions of the innovation conducted, was added in the post-test questionnaire (Appendix 5). This tool was mainly used to answer research sub-question 3 and triangulate and complement the video and interview data. This section aimed to examine the impact of the innovation on participants (seven questions), and participants' opinions of the thinking tasks conducted (six questions).

In order to facilitate answering, the pre- and post-test questionnaires were translated into the participants' native language, Chinese (see Appendix 5). A back-translation technique was not employed here because the researcher who composed the questionnaires and the co-workers were competent in both Chinese and English. The researcher did the first draft of the translation and it was then cross-checked by three of the researcher's colleagues. Any disagreement between the two versions was resolved through discussion and negotiation.

The pilot study was conducted in a college in Southern Taiwan. Some Chinese wordings and examples which were found inappropriate for the freshmen were modified based on informants' feedback.

Process

Questionnaires were distributed at both pre- and post-test by the researcher in the presence of the formal class lecturer. The pre-test questionnaire was distributed to both classes. In the post-test stage, the comparison class used the pre-test questionnaire version, while the innovation class was issued with the post-test questionnaire. This allowed comparison of data in Sections 1-3 and impact analysis of Section 4 in the innovation class.

Prior to commencing data collection, the researcher first explained the aims of the questionnaire. The content in each section was also briefly explained. Students were informed that it was an anonymous questionnaire and therefore answers could not be traced back to the individual respondent. Most important of all, respondents were assured that the results of the questionnaire would not affect their academic score. In addition, in order to avoid any social desirability bias (Bryman 2004), the researcher further explained that there were no right or wrong answers. The best answer was the true answer, expressing either positive or negative attitudes. Respondents were allowed to take their time to complete the questionnaire and to ask any questions they had about answering.

3.7.3 Interviews

Rationale

The rationale for using the interviews was to elicit more in-depth information from the participants. Quantitative data from questionnaires can provide a general picture of participants' ideas; however, the result often fails to offer an in-depth understanding of

the phenomenon under study. One way of avoiding the superficial understanding from quantitative data is to complement it with interviews. These data were used to answer research sub-question 3 and to triangulate and complement with video and questionnaire data.

There are three types of interviews: structured, semi-structured, and unstructured interviews (Berg 2006). The present study adopted a semi-structured interview as a means of data collection because as Richards (2003) argued, when the aim of the researcher is to explore particular lines of inquiry or to find out how different people view the same things, a semi-structured interview is more appropriate.

Design

Semi-structured interviews were conducted one to one with the six study group members and ten non-study group members in the innovation class after the post-test. In the study group of three, however, specific question was asked to individual in turn. This gave students the opportunity to listen, respond and also interject, but most of the time it was one to one conversation. Questions mainly related to participants' perceptions of and attitudes towards the impact of the HOTS approach on their thinking and speaking. Interviewees were asked to talk about six perspectives: a) whether they liked this teaching approach, b) whether there were more opportunities for speaking, c) whether there were more opportunities for thinking, d) whether there were any difficulties they encountered, e) whether there was any impact on their speaking and f) whether there was any impact on their thinking. These questions explored students' interest in this approach in a variety of different ways so as to provide in-depth and holistic understanding. It was also important to learn about how students considered this approach providing them speaking and thinking opportunities, the impact of these opportunities on their speaking and thinking performance, what difficulties they encountered and how they reacted to these difficulties.

Process

Interviews were conducted in Chinese and audio recorded. They took place in a quiet café for two reasons: it was hoped participants would feel more relaxed and therefore speak more freely in this environment, and it was somewhere students could sit down for a longer time without disturbance, which allowed the researcher to gain more time with each participant. The interviews were conducted with three participants at a time; they all took turns to answer the question paused so that everyone was given enough time to remember, conceptualize and verbalize. The voice recorder was placed on the table. Interviewees were informed that their opinions towards this innovation would not affect their academic score and their genuine views would contribute to the research. More in-depth information was elicited with these study group members.

Individual interviews with 10 non-study group members were conducted during the class break time. Each interview took around five to ten minutes and was taking place in the classroom. It was a snapshot of students' attitudes towards this innovation. The researcher with the voice recorder holding in the hand randomly approached any student who was free to talk. First of all, students were informed of their right to decline to be interviewed by the researcher, a choice made by two of the participants. Most of the participants, however, were willing to contribute their perceptions and ideas and they were encouraged to talk about their true opinions by informing them the interview would not affect their academic score (an extract from the transcription of an individual interview can be found in Appendix 6).

In the process of interviewing, the researcher tried to avoid asking leading questions (Oppenheim 1992). For example, the question 'Did you benefit from this innovation in terms of speaking? If yes, what was it?' was posed, instead of asking 'How did you benefit from this innovation in terms of speaking?'.

The mood of the informants and environmental variables which might affect the information given by informants (Nisbett and Wilson 1977) were taken into consideration. The researcher tried to reduce their anxiety and establish a rapport with participants (Oppenheim 1992) by telling them there was no right or wrong answers and any positive or negative feedback was welcome. While interviewees were talking, the researcher played a role as a listener and did not make any comments on their response, so as to avoid misleading or affecting interviewees' response.

Misunderstanding and ambiguity were minimised as much as possible. During the interview, it was found that one study group member misunderstood the idea of 'questions which are intellectually challenging', stated in the questionnaire (Section 3, no. 6); the researcher discovered that what she claimed in the interview was not consistent with what she stated in the questionnaire. This misunderstanding was clarified immediately, as argued by Legard et al. (2003) and Hughes (2002), concerning the advantage of interview, and a clearer picture of her perception of higher-order questions was shaped. Sometimes interviewees did not respond directly to the question, and the researcher needed to guide them to elaborate on the point explored. This is a considerable advantage for the researcher in conducting interviews: it is possible to obtain valid and more profound information.

3.8 Analysis

This section looks at the tools used for analyzing the data collected. Before discussion of these tools, it is important to have an overview of the process of data analysis and interpretation so as to fit with the criticisms around mixed-method design. Figure 10 demonstrates the framework for data analysis and data interpretation used in this study. Video recording and interview data were quantified using content analysis, in which students' utterances were converted into numerical codes that were represented statistically. Content analysis, according to Holsti (1969), is a technique which makes

inferences by objectively and systematically identifying specified characteristics. With regard to the questionnaire data, since the number of respondents was over thirty in each class, data were computed by using ANCOVA and Pair Samples T-test in the statistical program SPSS 15 (Statistical Package for the Social Sciences, 15 version). Data were triangulated and complemented at the analysis and interpretation stages, which is in line with Hammersley's definition of a mixed-method approach (1996).

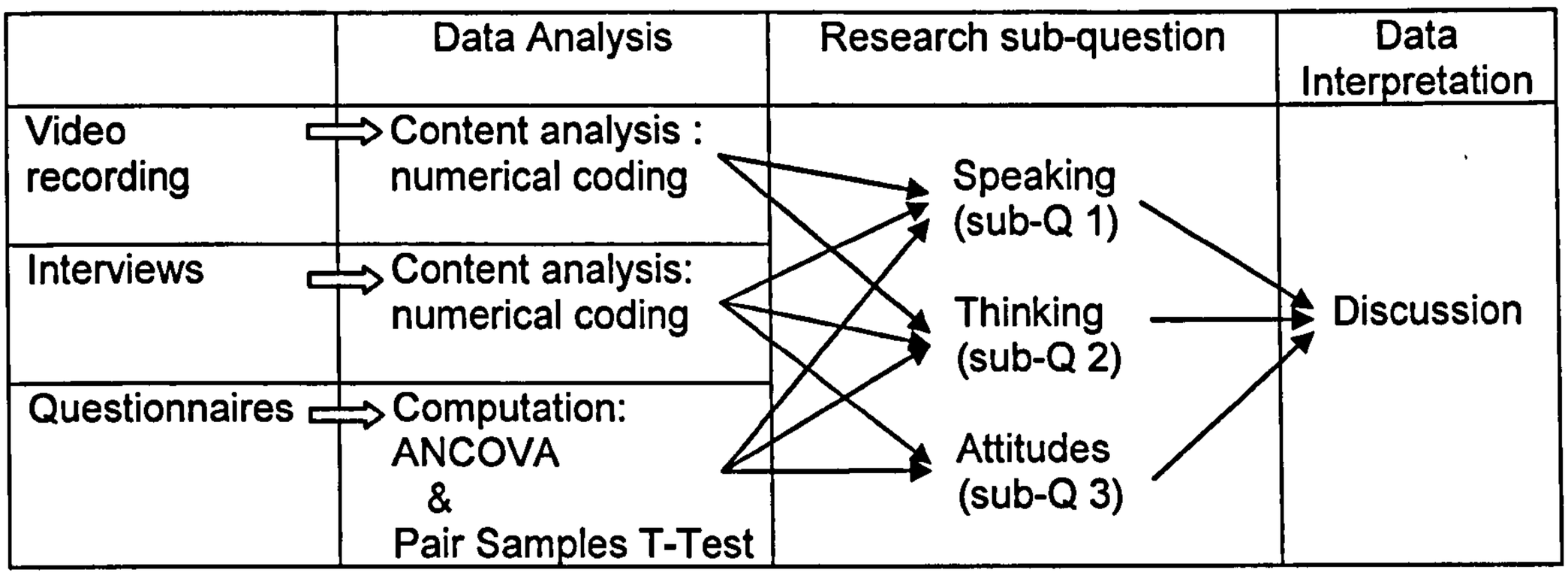


Figure 10. A framework of data analysis and interpretation

3.8.1 Video recording

Video data were analysed in terms of examining the length of utterance, grammatical complexity, the overall speaking proficiency, and the *difficulty* level of higher-order thinking skills. These data were collected from students' simultaneous discussion. The total recorded run length was different in each discussion, since some technical problems occurred during recording as discussed in Section 3.7.1. The shortest run length was 8 minutes. In order to make comparison easier, it was decided to use the first eight minutes of each transcript for analysis. Some excerpts extracted from the video data were provided in analysis chapter to exemplify the themes occurred from the interview data and analysing aspects of speaking and thinking performance.

The oral data were transcribed by the researcher and an English teacher (for the transcription system of utterance data, see Appendix 7); the researcher dealt with the

data from the innovation class while the English teacher handled that from the comparison class. The transcriptions were then swapped and cross-checked.

Before describing the analysis tools, it is necessary to note that, as far as Green and Harker (1988) were concerned, linguistic outcomes produced in a social context are complex and challenging to assess because the communicative performance of a student is contingent upon the actions of the other speakers. But, it was decided to analyse each individual's speaking performances. There were three reasons for this. First, working with the thinking tasks, individual students had space for monologue, in which they elaborated their own ideas before interacting with the others. Thus, the individual's speaking performance could be observed. Second, it was important to evaluate individual's coherence, a sign of communication and comprehension (Schiffrin 1994), when carrying out the thinking tasks in a social context. Coherence is also a criterion in the IELTS Test (Appendix 8, IELTS Speaking Band Descriptions) which assesses speaking proficiency. Third, examining the individual speaking performances revealed whether an overall improvement was limited to certain individuals or extended to the majority of students. As Pica and Doughty (1985) claim that when undertaking tasks in groups, certain group members may dominate the interaction while others fail to participate. Supporting Pica and Doughty, Bryman (2004) also argues that quantitative research tends to ignore individual performances and the knowledge produced may be too general for direct application to specific individuals; it may not be clear whether the improvement resulted from a specific treatment is made by certain students or is general.

When cognitive performance was examined, a general, holistic view was taken, not looking at individual's cognitive performances. The decision was reached because the high cognitive performance could be achieved collaboratively by the group members. For example, one higher-order thinking idea could be generated by two or three

students together. In addition, due to the time constraints, this study did not collect additional data for examining individual's cognitive performances.

One last important issue that needs to be noted is that the sample size of the study group members, six in each class, was too small to give reliable results to be analysed using SPSS. Thus, the analysis of video data will be presented based on descriptive statistics, such as mean scores.

Mean turn length

Mean turn length was adopted to examine the change in the length of students' utterances before and after the innovation. This was calculated using the total number of words divided by the total number of turns (Ellis 2005). In the present study the total number of words meant that all words were counted except fillers, repetition, the reading questions on the handout and Chinese words (see Appendix 8 for word count system). Ellis (ibid.) did not define the total number of words. The researcher e-mailed Ellis and inquired about the definition. He suggested that fillers and repetition should be excluded. With regard to excluding questions being read, there were two reasons for this decision: one was that the researcher considered the utterance of reading questions on the handout to be not authentically generated and the other was the amount of time spent reading questions was far too great in one specific group leading to confusion in the results. Compared with other study groups where one question had only been read once, the members in this particular group would read the question individually before he/she answered it, totalling three times for each question, thus the high frequency of reading questions led to an increase of the length.

Words and turns were tallied manually by the researcher because the measurement did not involve coding where bias might occur and the objects to be tallied were easily

identified. The results were demonstrated with bar charts generated using Microsoft Office Excel.

Grammatical complexity

Grammatical complexity was determined by measuring the amount of subordination, calculated by using the total number of separate clauses divided by the total number of AS units (Foster and Skehan 1996). The AS, the c-unit and T-unit can be considered as a basic unit when examining grammatical complexity in speaking. In the current study the AS unit was selected to be the basic unit instead of the T-unit or c-unit and this was justified as follows.

The T-Unit was defined by Hunt (1965, p. 20) as:

> 'a main clause with all subordinate clauses attached to it.'

Although this measure of subordination works well when analysing written texts, it works less well for oral discourse (Ellis 2005; Foster et al. 2000). Spoken data is not as tidy and clear cut as written data. Thus, the T-unit was not always easily identified with the complex features in spoken language, such as segmentation.

The c-unit was first defined by Loban (1966, p. 5 cited Foster et al. 2000) as:

> 'grammatical independent predication(s) ... answers to questions which lack only the repetition of the question elements to satisfy the criterion of independent predication....*Yes* can be admitted as a whole unit of communication when it is an answer to a question such as *Have you ever been sick*?'

Some researchers (e.g. Pica et al. 1989; Chaudron 1988) further defined the c-unit using Loban's definition. However, Foster et al. (2000, p. 361) argue that

> 'a major problem with these Loban-based definition is their seeming exclusion of elliptical constructions which arise *within* a speaker's turn rather than link to an interlocutor's question.'

The definition of a c-unit is not comprehensive in terms of covering different types of oral discourse and it does not provide detailed examples, which may lead to difficulties in applying it to real practice.

Foster et al. (2000, p. 365) define the AS unit as:

> 'a mainly syntactic unit...a single speaker's utterance consisting of an independent clause, or sub-clausal unit, together with any subordinate clauses(s) associated with either.'

Compared with the T-unit and c-unit, AS units are better in manipulating the segmentation of oral data and offer much clearer definitions in coping with the different degrees and types of embedding and attachment in the longer utterance (see Foster et al. (2000) for a demonstration of how to deal with different types of utterance). Thus, the AS unit was chosen to be the basic unit in this research.

In the process of coding the oral data, however, the researcher discovered that the AS unit did not define how to code an incomplete sentence. Incomplete sentences occurred while the speaker was still at the planning stage of producing ideas, and without finishing the sentence the turn was taken by another speaker. The researcher identified this kind of incomplete utterance and considered it as a AS unit (for AS units and clauses coding system and an example of the coding, see Appendix 8).

The rationale for using a predetermined template for AS units and clauses rather than categories derived from the text itself was that AS units and clauses are identified as the language units, such as words, phrase and utterance, and a predetermined

template is applied when analysing language units (Crabtree and Miller 1999). The procedure of organizing and coding the data with a predetermined template is called 'template', which is further explained in Appendix 8, organizing style of analysis.

To control for bias, the coding of AS units and clauses was carried out by the researcher and a co-rater who is an English teacher. The researcher first explained the concepts of AS units and clauses to this English teacher and went through the examples of coding AS units and clauses given in Foster at al. (2000). The two raters discussed the concept, then reviewed and negotiated the coding.

To validate the equality of coding between two coders, inter-rater reliability (Silverman 1993) was tested. To establish adequate inter-rater reliability for this study, a piece of transcription of students' utterance data was evaluated by the two raters. Using Miles and Huberman's (1994) inter-rater reliability formula (reliability = number of agreements / (total number of agreements + disagreements)), the inter-rater reliability rate for AS units coding was determined at 95 % agreement. Of the total 107 AS unit coding decisions made, there were only 5 differences. The inter-rater reliability rate for clause coding was determined at 92% agreement: of the total 126 coding decisions made, there were 10 differences. Based on these two ratings, both the accuracy and reliability of using these coding instruments, AS units and clauses, met the general check-coding standard, which is required to be in the 90% range (Miles and Huberman 1994).

After the consistency had been established, the two raters started to code the AS units and clauses. Any differences which occurred in the process of coding were discussed and negotiated by the two raters through mutual conversation. Intra-rater reliability was tested three months later by the researcher herself and was determined at 91% agreement for AS units and 95% agreement for clauses.

Having completed the coding of the AS units and clauses, to facilitate the calculation of the total number of AS units and clauses the researcher transferred the AS units and clauses to Excel. The results were then presented using bar charts.

Speaking proficiency

Students' overall speaking proficiency was rated by a native-speaking test expert who had done the work for more than 10 years and was familiar with the IELTS speaking band descriptors. The public version of IELTS speaking band descriptors (Appendix 8) was used to provide criteria when evaluating the overall speaking proficiency. The score ranged from 0-9. Scores given in this study was based on .1 scale instead of .5 scale, e.g. 5.1, 5.2, which shows the change of students' utterance more subtly. In the process of rating, the researcher played the audio recording for the rater in a non-sequential order of pre-, post-, and delayed post-test in order to diminish rater's bias of expecting certain outcomes. However, there was one limitation: the standardization of rating. The data were rated separately in four weeks, one study group per week. The standardization between the four groups may have been slightly different when rating. The results of the speaking proficiency were also examined to see whether there was an extreme value by using box-plot in SPSS.

The frequency of HOTS

Cognitive performance was investigated by examining the difficulty level of HOTS. The difficulty level of HOTS was determined by two factors in this study: the frequency of HOTS and the function of HOTS. The former was analysed using the six thinking levels of Bloom's taxonomy (Section 2.4), while the latter was evaluated with three categories: answering the questions listed on the handout, answering others' enquiries, commenting or building on others' opinions.

The analysing tool, the six thinking categories, is first explained. After applying the six thinking categories to the text, it was found that some utterances, for example questions, were not able to fit within these six types of thinking categories. Therefore, those units which did not fit were further categorized as higher-order and lower-order questions (Section 2.5). This type of organizing style of analysis is called 'editing' (Appendix 8, organizing style of analysis), as stated in Crabtree and Miller (1999). Editing means that the interpreter engages with the text without a predetermined template and searches for meaningful units. The units both stand on their own and relate to the purpose of the study. The identified units are then sorted into categories (for coding system of six thinking levels, see Appendix 9).

The process of categorizing six thinking levels was carried out manually by the researcher and a co-coder instead of using Nvivo 8. Nvivo 8 allows researchers to code the text and to retrieve all those sequences of coded text, which means the computer can take over the manual tasks, like copying and pasting the related text together. However, the categorization of thinking skills was strongly context dependent. For example, an utterance of higher-order thinking idea which occurred the first time was categorized as higher-order thinking; yet, if the same higher-order thinking idea occurred again a second time, it was then categorized as lower-order thinking (a rationale for this, see Section 2.4.2). Due to this coding principle which underpinned the system, the researcher decided to analyse the data manually. Inter-rater reliability was built up before coding, which was determined at 80% agreement. The agreement was slightly lower than the general check-coding standard, 90% range (Miles and Huberman 1994). It is hoped that the reliability was enhanced through the discussion and negotiation of the differences occurred in the process of coding by the two raters. The intra-rater reliability was tested three months later by the researcher herself and reached 84% agreement. Having gained this slightly lower agreement, the

researcher had gone through the whole data again to identify the disagreement and make certain necessary changes for the sake of consistency.

The function of HOTS

Regarding the function of HOTS, an analysing tool was developed by the researcher. After looking at the whole text, the function of HOTS were categorised into three groups: answering the questions listed on the handout, answering others' inquiries, and commenting or building on others' opinions (for the coding system of the function of HOTS, see Appendix 9). The purpose of the idea generated using HOTS was examined according to these three categories. The organizing style of analysis was 'editing' (Crabtree and Miller 1999).

This analysing tool was developed to examine how students were engaging with higher-order thinking skills in a social context. The cognitive workload is heavier in answering others' inquiries and commenting on others' opinions than answering questions listed on a handout due to the effort needed to analyse the information received and to justify what information can be commented on. The HOTS are used both during information processing and generating a response, and this requires much greater effort than simply answering questions on the handout. As Sousa (2001) stresses, the difficulty level enhances, while the amount of effort expended increases. Consequently, while the frequency of listening to others' inquiries and making comments rises, the difficulty level of HOTS also increases.

The coding process was carried out by the researcher and the co-coder. The inter-rater reliability reached 100% agreement. Two months after the coding, the intra-rater reliability was tested by the researcher herself and reached 95% agreement.

3.8.2 Interviews

The interview data were transcribed and analyzed in Chinese, thus avoiding a considerable amount of translation work and saving time. The organizing style of analysis applied the style of editing (Crabtree and Miller 1999). After going through the whole text, the researcher identified three main categories, namely the attitudes towards this innovation, the opportunities for thinking and talking, and the perceptions of the effect of the innovation on students' performance. Within these three main categories, the units were then again identified and subcategorized. The coding was completed by the researcher and an English teacher who contributed to the work of cross-checking the categorization. Different opinions regarding the categorization were discussed and resolved.

The results of interview data were translated into English after the analysis was done. The translation was then carefully checked by an expert bilingual in both Chinese and English.

3.8.3 Questionnaires

Questionnaire Sections 1-3 were analyzed using the statistical program SPSS 15. ANCOVA (Analysis of Covariance) was employed to examine the effect of the innovation on students' attitudes and perceptions. It compares the means of two or more samples while controlling for the variation due to an extraneous variable (Tashakkori and Teddlie 1998, p. 116); the extraneous variable in the present study was the pre-test. In other words, if any variables (covariates) are known to influence the dependent variable being measured other than the independent variable, then ANCOVA is ideally suited to remove the bias of these covariates (Field 2000). Take the present study as an example, the result of pre-test is one variable other than the independent variable, the treatment, which could affect the outcome, the dependent variable. To examine whether there is a significant difference of covariate at pre-test

between the two classes, homogeneity of with-in regression of One-way ANCOVA is employed. If there is no significant difference of covariate, the analysis of One-way ANCOVA is then applied to continuously examine the effect of the innovation conducted.

In addition, students' attitudes within one class before and after the treatment were also compared using Paired Samples T-Test. Peers (1996) points out that when two group means are compared from the same subjects, that is, paired measurements for each subject in the sample, Paired Samples T-Test should be considered.

Both ANCOVA and Paired-Samples T Test were analyzed based on the categories listed in Table 3. The first three thinking levels in Bloom's taxonomy - knowledge, comprehension, and application - were categorized as lower-order thinking, whilst the last three thinking levels, namely analysis, synthesis, and evaluation, were categorized as higher-order thinking. Analyzing categories were: a) the importance of using lower-order thinking in L2 classrooms, b) the importance of using higher-order thinking in L2 classrooms, c) the ability of answering lower-order questions, d) the ability of answering higher-order questions, and e) questioning-answering behavior.

Table 3. The categories of questionnaires

Section	Questions	Categories
Section 1	Q1-9	a) The importance of using lower-order thinking in L2 classrooms
	Q10-18	b) The importance of using higher-order thinking in L2 classrooms
Section 2	Q1-9	c) The ability of answering lower-order questions
	Q10-18	d) The ability of answering higher-order questions
Section 3	Q1-9	e) Questioning-answering behaviour

Questionnaire section 4 was analysed using descriptive analysis and the percentage of students responding to the questions presented.

3.9 Reliability and validity

Issues relating to reliability and validity were taken into account when designing the research, collecting and analyzing and interpreting the data. Quantitative and qualitative research methods have different levels of reliability and validity. This section discusses the different aspects of reliability and validity and how they relate to the current research. Both Oppenheim (1992) and Cohen et al. (2007) stress that reliability is an essential precondition of validity, so reliability will be first examined.

3.9.1 Reliability

According to Bryman (2004, p. 28) 'reliability is concerned with the question of whether the results of a study are repeatable'. That is to say, reliability refers to the extent to which the findings can be replicated if the procedures and processes are repeated with the same respondents. The goal of reliability is to minimize the biases in a study (Yin 2003).

Internal consistency estimates reliability by grouping questions in a questionnaire that measures the same concept. Before collecting questionnaire data, internal reliability of the questionnaires was considered by employing Cronbach's Alpha test (Peers 1996) and reached the satisfactory level of .94. The relationship between the researcher and the respondents had been considered while collecting the questionnaire data. The results of the questionnaires collected from the post-test stage may have been affected as a result of the close relationship between the researcher and the innovation students.

A high reliability of interview data collection could be assured by considering the informants' mood, building a rapport and avoiding leading questions. However, the reliability may have decreased as a result of the bias caused by the characteristics of the interviewer and interviewees.

Instrument reliability was ensured by looking at the equivalence of the tests conducted. For example, the type of tasks, the topics and the number of questions were taken into consideration when designing the tests. While analyzing the video data, including grammatical complexity and the function of HOTS, inter-rater and intra-rater reliability were assessed and passed the general check-coding standard.

Yet, it needs to be recognized that the inter-rater and intra-rater reliability of six thinking levels coding did not pass the general check-coding standard; they were slightly lower than the requested standard. Therefore, the reliability of this finding needs to be considered when reading results. However, it is hoped that the reliability was enhanced through the discussion and negotiation of the differences occurred in the process of coding by the two raters.

In addition, the reliability of speaking proficiency ratings can be questioned due to two reasons. One was that there was only one rater, so the bias carried by the rater was not controlled. The other was the rater did not standardize the rating between each assessment. The reliability of the interview data coding should also be considered due to a lack of a co-coder. It is hoped that the reliability improved through the work of cross-checking the categorization done by a Taiwanese teacher of English language.

Students' reactive effects may have occurred resulting from the presence of the voice recorder and camcorder and may thus have affected the reliability of the study. Nevertheless, it is hoped that this issue may have been minimized in this study by the

fact that the voice recorder was used throughout the innovation and participants were accustomed to it.

On the whole, data were collected mainly by the researcher. Thus the reliability may have been affected by researcher bias (Fine et al. 2009). However, in the process of collecting questionnaire and interview data, some essential techniques to avoid bias had been implemented as discussed in Section 3.7.2 and 3.7.3.

Finally, the conceptual framework, the research design, thinking task design, lesson plan, participants, implementation of the innovation and data collection procedure were stated crystal clearly for a replication. When the innovation results were highly reproducible, the reliability of the study increased.

3.9.2 Validity

Validity is concerned with the accuracy and appropriateness of the different methods chosen to answer the research questions. According to Bryman (2004, p. 28),

> 'validity is concerned with the integrity of the conclusions that are generated from a piece of research.'

Different research tools have different levels of validity. The use of triangulation and complementation in a mixed-method approach can strengthen the validity of each data collection tool, and thus enhance the validity of the study overall. Bryman (ibid., p. 28) distinguishes four main types of validity:

> *Measurement validity*: the extent to which a chosen measure of a concept reflects the concept which it is supposed to;
> *Internal validity*: relates to the issue of causality;
> *External validity*: is concerned with generalization of the results of a study;

Ecological validity: is concerned with whether the findings are applicable to natural social settings.

Regarding measurement validity, this is pertinent to the questionnaires. Whether these questionnaires really measured the extent to which students reflected on their ability when using different thinking levels; that is, to what extent students could clearly distinguish which level of thinking they were using and justified it. Nevertheless, it is hoped that the validity was enhanced through providing examples for each thinking level.

The analysis tool used to examine the function of HOTS by exploring whether a higher-order thinking idea elaborated was used for answering the question on the handout, answering enquiry or commenting on others' view in a social context, was developed by the researcher; it has not been validated in the literature. Therefore, its validity needed to be considered.

As to internal validity, there are three key issues. First, study group members in both classes where the video data were collected, were comparable before the innovation. They were all non-English majors from the same university, and they used the same textbook and shared the same English lecturer with the same pedagogy in the first semester. The average score of English language for the first semester in the two groups were approximately the same. In addition, none of them took extra English lessons after school. Their background arguably made them comparable in the context under evaluation. Yet, the different sampling techniques for choosing study group members in two classes could be a potential threat to the validity of the study. Study group members in the innovation class volunteered to take part in the study while those in the comparison class were nominated. This could imply that study group members in the innovation class might have higher learning motivation than those being nominated in the comparison class. Second, the researcher administered the

same measuring instrument for video data in all three tests in which the same coding systems were applied to gauge students' utterances for a certain aspect, e.g. grammatical complexity, the length of utterance, etc. Third, the study was a single case study design, yet one comparison group was included. All of these resulted in an enhancement of the internal validity.

There is, however, one potential threat to the internal validity regarding participants recruited: they majored in different subjects where the comparison class majored in maths, while the innovation class majored in Visual-Art Design. Their thinking styles can be different, thus affect the validity of this study.

The external validity of a relatively small scale innovation with only non-English majors as a sample is open to criticism. This study aimed at theoretical generalization (Yin 2003) and the sample was expected to include two types of L2 learners: non-English majors and English majors. Compared with the number of non-English majors, English majors is rather a small group among L2 learners in Taiwan. Yet, without examining the group of English majors, the theoretical generalizability could decrease.

The embedded design of this case study is another potential threat to the external validity. The six study group members, as the embedded units, were the focus where the video data were collected for examining the effectiveness of students' speaking and thinking performance. This small number could be a pitfall of the study design.

The study may have relatively low ecological validity due to the innovator effect. The innovation treatment might be ineffective because of the innovator. Lack of awareness about the ways in which thinking skills can be conceptualized and the implications for lesson design hinders teachers to carry out such thinking approach (Waters 2006). Conducting higher-order questions and designing the thinking tasks require a more

professional knowledge like an understanding of what HOTS are, how these skills can be integrated into L2 learning and the theories underpinned. It also requires teaching experience so as to carry out the implementation specifically; it can be difficult for a novice teacher to carry out such an implementation without the basic experience of teaching four skills (see Section 6.3 for further discussion).

3.10 Ethical issues

> 'Ethics is a matter of principled sensitivity to the rights of others' (Gilbert 2008, p. 45).

Being ethical confines the choices researchers can make in the pursuit of truth. Social science researchers must take into account the ethical effects of data collection from people and about people (Cohen et al. 2007). Ethical issues relating to the different data collection methods employed in the current study are addressed and discussed in this section.

In innovation research, participants should not be deceived and prevented from knowing the purpose and conditions of the research (Cresswell, 2003). Prior to commencing the study, this issue was taken into consideration. First of all, the director of the English Department was informed of the purpose for conducting the research and she orally replied with a warm welcome and encouragement. Then, students were also informed of the purpose of the study with a letter of consent stating the following: a) their right to participate voluntarily and withdraw at any stage, b) the issue of beneficence, c) a guarantee of no harm to them or disruption of their schedule, d) a guarantee of confidentiality, anonymity and non-traceability in the research, e) a guarantee that data would be destroyed within two years.

Cook (1986) argues that the ethical concerns in the second language learning domain are not as serious as those in first language learning experiments. The purpose for carrying out experimental teaching in the second language learning domain is to improve language teaching and learning. It is the degree of effectiveness and appropriateness that matters, and there is no real harm to the participants. Besides, although the video data contained personal images, and it was strongly connected to personal privacy, this study did not pry into personal affairs. The participants were assured that the data would only be kept by the researcher, viewed by people joining in the process of analysis and the researchers' supervisors and used for academic purposes. Therefore, no permanent damage to participants could result.

One ethical issue about observation relates to whether participants were prepared to be observed. This issue had been tackled before selecting study group members through oral inquiry. They all agreed to be video-taped.

Participants who learned better through a Grammar-Translation Method might not have benefited from this innovation due to different learning styles. For those who disliked thinking actively in the English class, the innovation involved in this study may have been a difficult experience for them. This implies that the balance between student need and pedagogy should be reached while conducting the HOTS approach.

A final ethical issue which needed to be considered is that participants who were not included in the study groups may have felt undervalued in some way. One study group member in the innovation class told the researcher that other participants thought that study group members were more important and played a crucial role, since the researcher had more contact with them, e.g. arranging time for interviews. In order to clarify the misunderstanding, the researcher declared again at the end of the semester

that all of their contributions, including study group members and non-study group members, were of equal importance and value.

3.11 Summary

This chapter has looked at the methodology underpinned this study in which the rationales for adopting pragmatist as a stand point, using a mixed-method approach and applying a comparative case study design were articulated. Participants recruited were university non-English majored freshmen and were chosen through non-probability sampling. To answer the research questions, data were collected through video, questionnaires and interviews and analysed and discussed based on the notion of a mixed-method approach. The reliability and validity of this study were justified and ethical issues of conducting the research were also tackled.

CHAPTER 4 - RESULTS

4.1 Introduction

In Chapter 3 (Section 3.3), it was argued that to gain a holistic and comprehensive understanding of using higher-order questions in L2 classrooms, it was important to examine not only the effect on speaking and thinking but also students' attitudes and perceptions. This chapter, first of all, presents the effects of the HOTS approach on students' speaking and thinking. Findings from the comparison data are compared to determine the efficiency of the output. Following this, perceptions of thinking skills and attitudes towards this innovation are demonstrated.

As stated in the methodology chapter, one advantage of a mixed-method approach is ensuring that quantitative and qualitative data complement and validate each other to provide overall coherence to the study (Section 3.5.1). Thus, based on the research questions, the results gained from the different data collection tools will be collated, to triangulate and complement each other at the analysis stage. The framework of data analysis was shown in Figure 10 (Section 3.8). Video data were the main data used to examine the effect on students' thinking and speaking, while interview data were the main data used to investigate students' attitudes and perceptions, as shown in Table 4. However, the interview and questionnaire data were also triangulate with the video data. Comparisons between the innovation and the comparison classes will be made to determine the effect. In these cases data from the innovation class will be presented first, followed by that of the comparison class.

Table 4. The priority of data used for analysis

Research questions	Video	Interview	Questionnaire
Sub-question 1: Speaking	1st	2nd	3rd
Sub-question 2: Thinking	1st	2nd	3rd
Sub-question 3: Attitudes & perceptions		1st	2nd

1st refers to the main data used for looking at the specific perspectives.
2nd and 3rd refer to minor data used for complementing and triangulating.

4.2 The effect on speaking

This section presents the results relevant to speaking. The results include students' perceptions of the speaking opportunities provided, perceptions of the effect on speaking, and quantitative evidence from the student talk in classrooms as recorded using video. This will show any effect on speaking performance which will triangulate with the perceptions.

4.2.1 Perceptions of opportunities for speaking

Four themes were found in the interview data on why this innovation offered more chances to talk: the nature of the thinking tasks, students' perceptions of thinking, extension of the topic and topic familiarity. These aspects will be discussed in turn. Questionnaire data will complement these data by providing a total figure.

The first reason is that the thinking tasks were designed around providing chances for speaking, such as asking students' opinions, requiring students to comment on one another's thoughts and to come up with a final decision within a group. As student L8 stated, the nature of the thinking tasks was perceived by students as enabling them to talk due to the need for discussion:

> L8: 'I think that the increase of opportunities for talking is due to the demand for discussion.'

Thinking is the second reason which allowed students to talk more. Thinking here meant higher-order thinking. Conducting higher-order thinking skills (HOTS) allowed students to come up with the reasons or ideas so that they could talk. They thought, so that they could talk, as L10 reported:

> L10: 'If I do not think, I would have nothing to talk about.'

Thinking enabled students to extend the topic. Extension of the topic emerged as another reason for students talking more. The thinking tasks allowed learners to control and extend the topic of interest, thus opportunities for speaking were also created. This meant that the thinking tasks provided speaking opportunities, yet it was *thinking* which further generated speaking opportunities. As L9 said:

> L9: 'I can further explore and elaborate anything which is not discussed by my group members.'

To conduct thinking, topic familiarity was essential and this was the last reason for increased talk. Students' schemata were the foundation for the discussion, which helped them to activate their HOTS. Insufficient background knowledge could impede learners' thinking. One student, L12, argued that sufficient background knowledge together with interest towards the topic enabled him to think and elaborate more.

> L12: 'If I am not familiar with the topic, I am not able to think of anything and it becomes difficult for me to discuss with others...I am interested in dogs, so I talked a lot on this topic.'

The total number of students who perceived this innovation as offering speaking opportunities was provided by post-test questionnaire data (Section 3, No. 8), which showed a figure of 80%. The results indicate that the majority of students were aware

of the increase in chances for talk provided and the elements which enabled them to talk more. So it shows that the HOTS approach could provide opportunities for talk.

4.2.2 Perceptions of the effect on speaking

In addition to the speaking opportunities provided, interviewees also perceived a change in their speaking in terms of the amount of talk, fluency, use of a wider range of vocabulary items, substitution and extension. Examples of substitution and extension are provided to show how these speaking techniques were talked about by the students.

One aspect of the innovation was the enhancement of speaking fluency perceived by students. A number of respondents clearly felt that they could talk more fluently, which meant that they were getting better at retrieving and processing lexical items, sentence structures and phonological features, and delivering their thoughts more smoothly. Six of the sixteen interviewees agreed that their speaking fluency had improved. Although the number of students was not high, it indicates a greater opportunity for students to talk fluently was provided by this innovation, as L7 claimed:

> L7: 'The fluency really improves...That means I can speak more fluently.'

Some students may not necessarily increase their fluency, but they perceived they *talked more*, which meant the length of utterance became longer. Questionnaire data (Section 4, No. 1) revealed that the number of students who considered they talked more as a result of the innovation was 87.5% of the total, as L21 stated:

> L21: 'I do not clearly know how much I improve, but I am sure I talk more.'

Three of the sixteen interviewees claimed that due to the need to express their thoughts, they started to use vocabulary items which they had rarely used previously. The number of students reported it was low. However, it shows that this innovation stretched the speakers in terms of lexical processing. Interviewees, as claimed by L13, started to search for previously learned items and to use them in speaking.

> L13: '[Now I often use words which] I wouldn't use before.'

Interviewees also discovered that they developed the use of substitution as a technique for speaking. Nine of the sixteen interviewees reported that they were able to use substitute words or paraphrase when unable to find the vocabulary item they wanted. As L19 and L20 stated:

> L19: "I can use other words to substitute. For example, I don't know the word '*humble*', I would explain it as *don't tell others I am good.*"
>
> L20: "When we played 'Guess What I Say', I used alternative words to explain without mentioning the idiom."

Another speaking technique which students saw themselves developing was extending the topic. With the privilege of controlling the topic, students learned to exercise this technique. Five interviewees clearly perceived that their ability to extend the topic was enhanced, as reported below by L9. Although there were only five students reported it, this finding shows there was potential for students to develop such a skill with the HOTS approach. To allow such a skill to be developed might take a longer time through practice.

> L9: 'I can extend the topic from a certain point.'

The following provides two excerpts where students use substitute words and extend the talk, extracted from the video data. These provide evidence of the improvements perceived by students in the learning process. Excerpt 1 demonstrates the technique of substitution. L9 inquired about the word '*reliable*'. L7 used the other word '*trust*' embedded in a sentence to explain what she meant by '*reliable.*'

Excerpt 1- 'The use of substitute words'

Line	
1	L7: ... I think good friend should be easy going, just like Amy and Ken...and faithful
2	just like Ken, and reliable
3	L9: *reliable*?
4	L7: ...*the good friend should trust me and I have to trust her*...

Excerpt 2 gives an example of a student extending the topic: the technique of extension. L8 claimed that her ideal mate should be responsible, so he would come home and hug the family even though he did not love her. L9 extended L8's idea by saying '*you can trust him.*' L8 then further built on this opinion and gave an example in line 6-7, saying that she would trust him not to have an affair with other girls.

Excerpt 2 - 'Extending the topic'

Line	
1	L8: my ideal mate have to responsible, his character maybe have a lot of sense of
2	responsibility, when we old and he didn't love me, it's ok, he will go home and
3	hug us...
4	L7: ok, I know
5	L9: ya, *you can trust him*
6	L8: *ah, yes, yes, and I think his responsibility may let him do everything good, not to*
7	*out of control, to have a date with other girl, because he have a family...*

These findings triangulate with the students' perceptions of the changes in their utterances and speaking techniques and therefore add confidence in these results. It indicates that students were aware of the improvement of speaking in the learning process and what they perceived did reflect what was observed in the classroom.

4.2.3 The effect on speaking

This section presents evidence of the impact on linguistic competence and how good students could command the spoken language in a social context. The video data

were used to analyse the length of utterance, grammatical complexity and the overall speaking proficiency. The findings demonstrate the progress of speaking made by the innovation class and then compared to the comparison class.

Length of utterance

As previously outlined (Section 3.8.1), Mean Turn Length (MTL) was determined by using the total number of words divided by the total number of turns (total words/total turns). Total number of words included all words except fillers, repetition, the reading questions on the handout and Chinese words. The numbers of these excluded features were also tallied (for results, see Appendix 10).

Figure 11 shows the results of MTL. The MTL figure in the innovation class improved only slightly at post-test, but increased considerably at delayed post-test, from 7.7 to 11. This indicates that the length of utterance improved steadily from pre-test to post-test and the ability to produce longer utterances was retained at delayed post-test. On the other hand, the MTL of the comparison class improved considerably at post-test, from 2.57 to 6.56 words, and dropped slightly at delayed post-test.

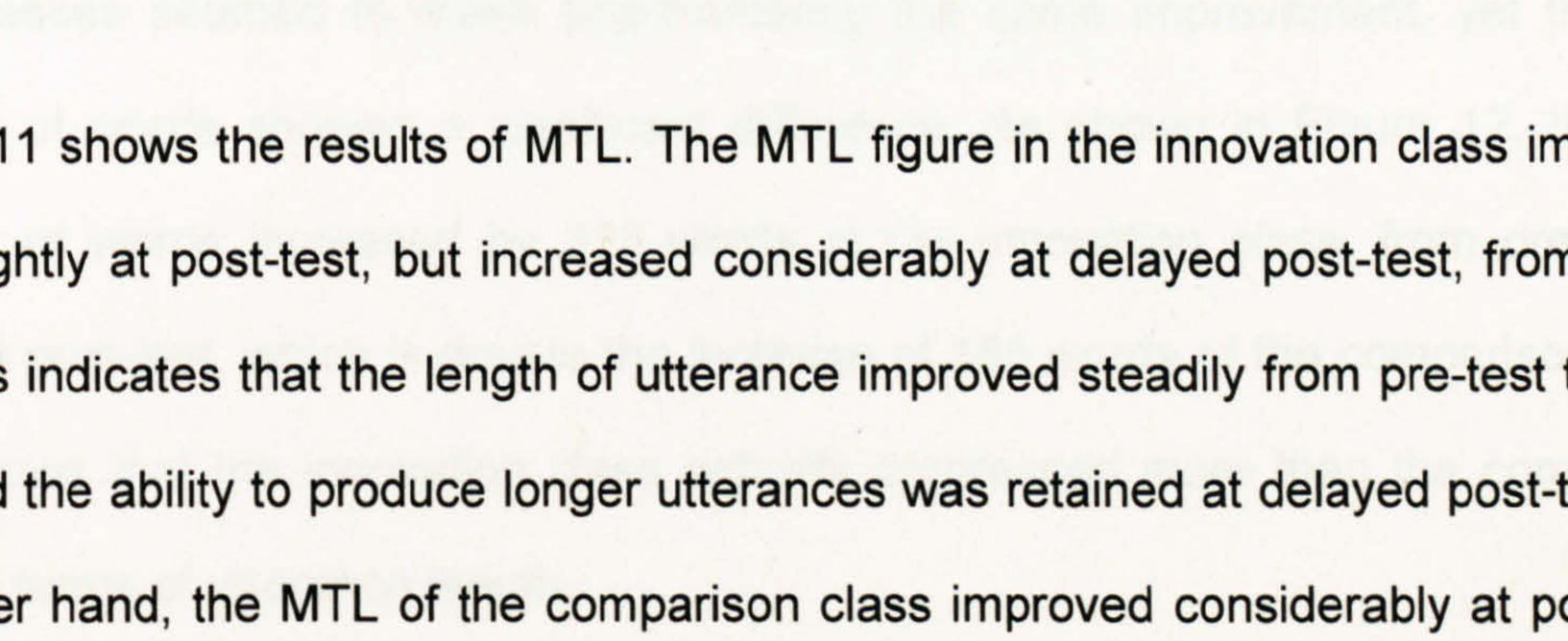


Figure 11. Mean turn length

The increase in length of the utterance in the innovation class, showed more improvement at delayed post-test, whereas, only a slight improvement was exhibited

at post-test. This may be explained by the fact that at post-test group members had different interpretations of certain words, e.g. 'faithful', 'loyal', 'reliable', and they got stuck in the discussion (see Section 4.3.2 Excerpt 4 for an example). They may have devoted too much time to thinking about how to clarify the different interpretations of these words, which meant that total number of words did not increase significantly. At delayed post-test, the students did not encounter this problem; as a result, the total number of words increased greatly. This implies that students had made improvement during the innovation; however, this improvement was more significant at delayed post-test.

Both classes seemed to make approximately the same improvement, yet the total number of words showed a significant difference. As shown in Figure 12, the total number of words increased by 313 words in the innovation class, from pre-test to delayed post-test, which is double the increase of 155 words of the comparison class. This shows that the innovation class actually progressed more than the comparison class in terms of utterance length.

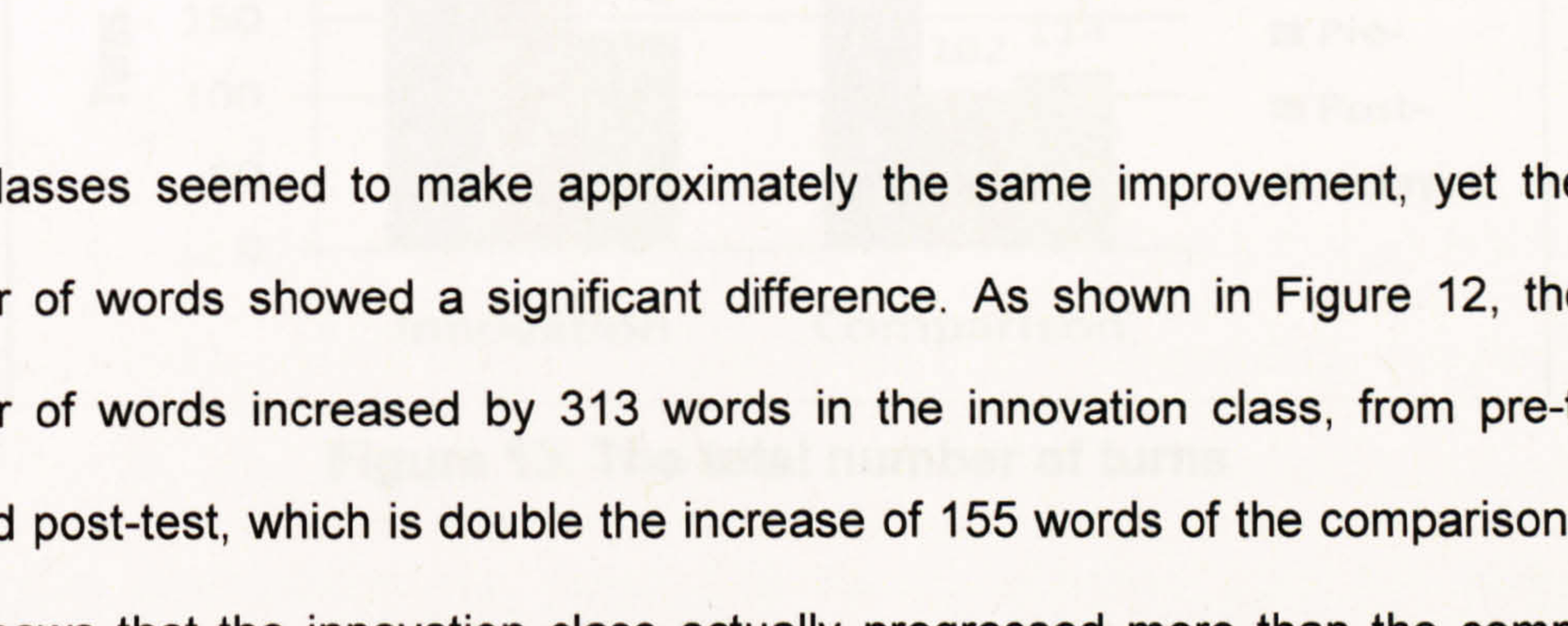


Figure 12. The total number of words

It is interesting to find that, as shown in Figure 13, the total number of turns in the comparison class decreased by around half at post-test and delayed post-test. In contrast, the innovation class dropped only slightly. This affected the results gained.

The more the number of turns drops, the higher the MTL figure gains. These results indicate that the MTL figure increased in the comparison class mainly due to the dramatic drop in the total number of turns, rather than an increase in the total number of words. In addition, the significant fall in number of turns in the comparison class indicates a decrease of group interaction.

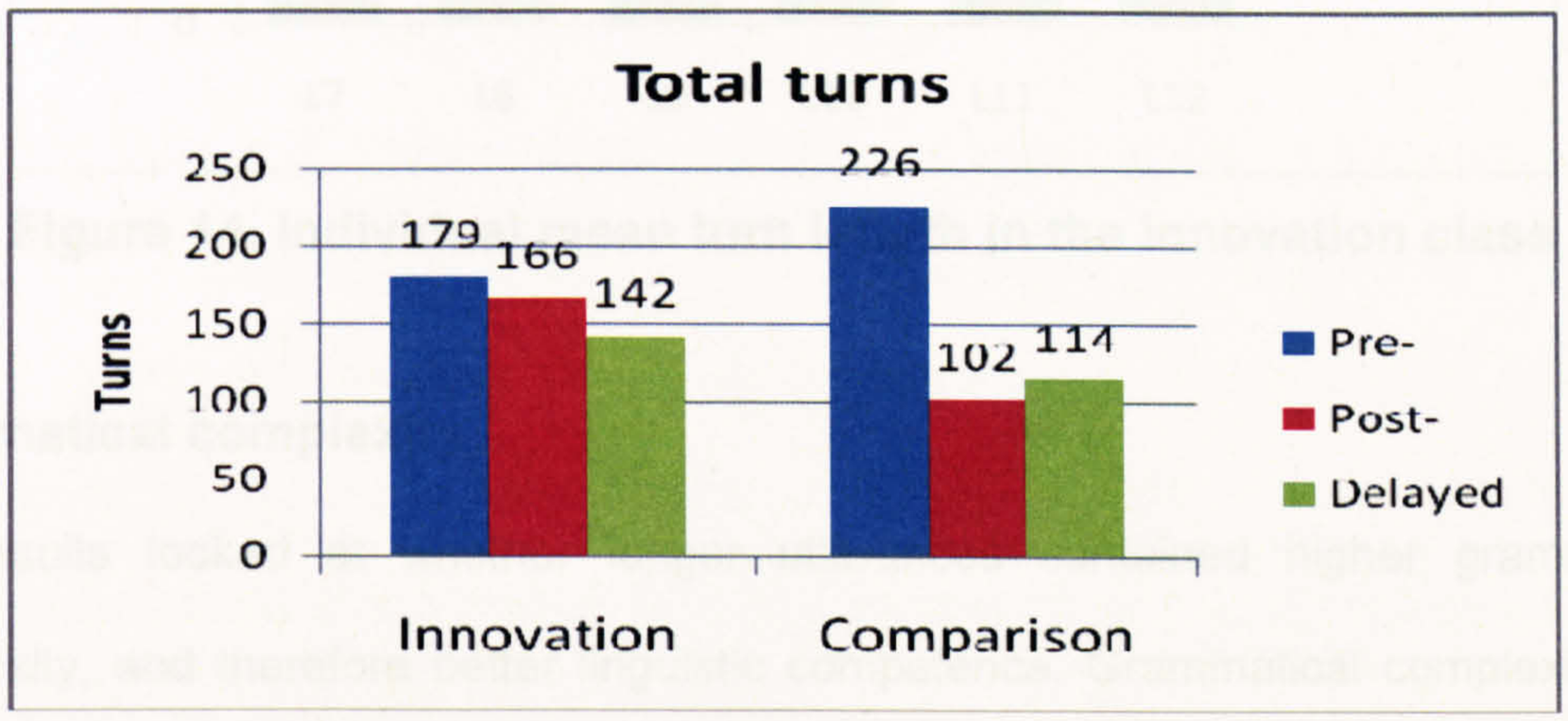


Figure 13. The total number of turns

The above section showed an improvement of the length of utterance in both classes in that the innovation class had a more significant outcome. The individual number of students in the innovation class making an improvement in MTL was further examined. Figure 14 demonstrates that in the innovation class there were only two out of six students (L10, L12) who improved at post-test stage, which shows that the improvement made was uneven. However, five students, all except L10, made an improvement from post-test to delayed post-test stage, potentially indicating a generic progress.

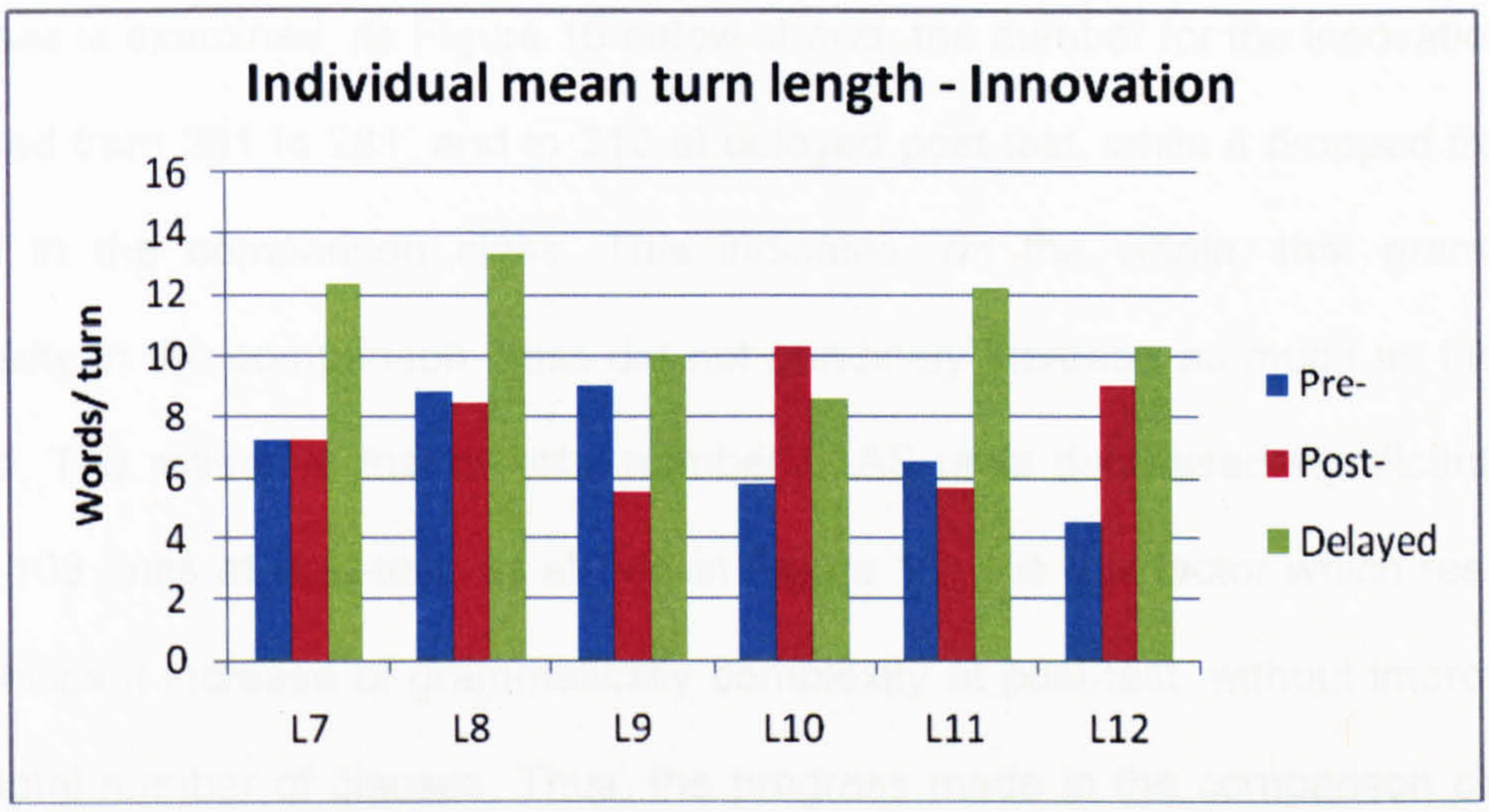


Figure 14. Individual mean turn length in the innovation class

Grammatical complexity

The results looked at whether longer utterances contained higher grammatical complexity, and therefore better linguistic competence. Grammatical complexity was determined by using the total number of clauses divided by the total number of AS units (total clauses/total AS units). The results obtained from the analysis of grammatical complexity are shown in Figure 15.

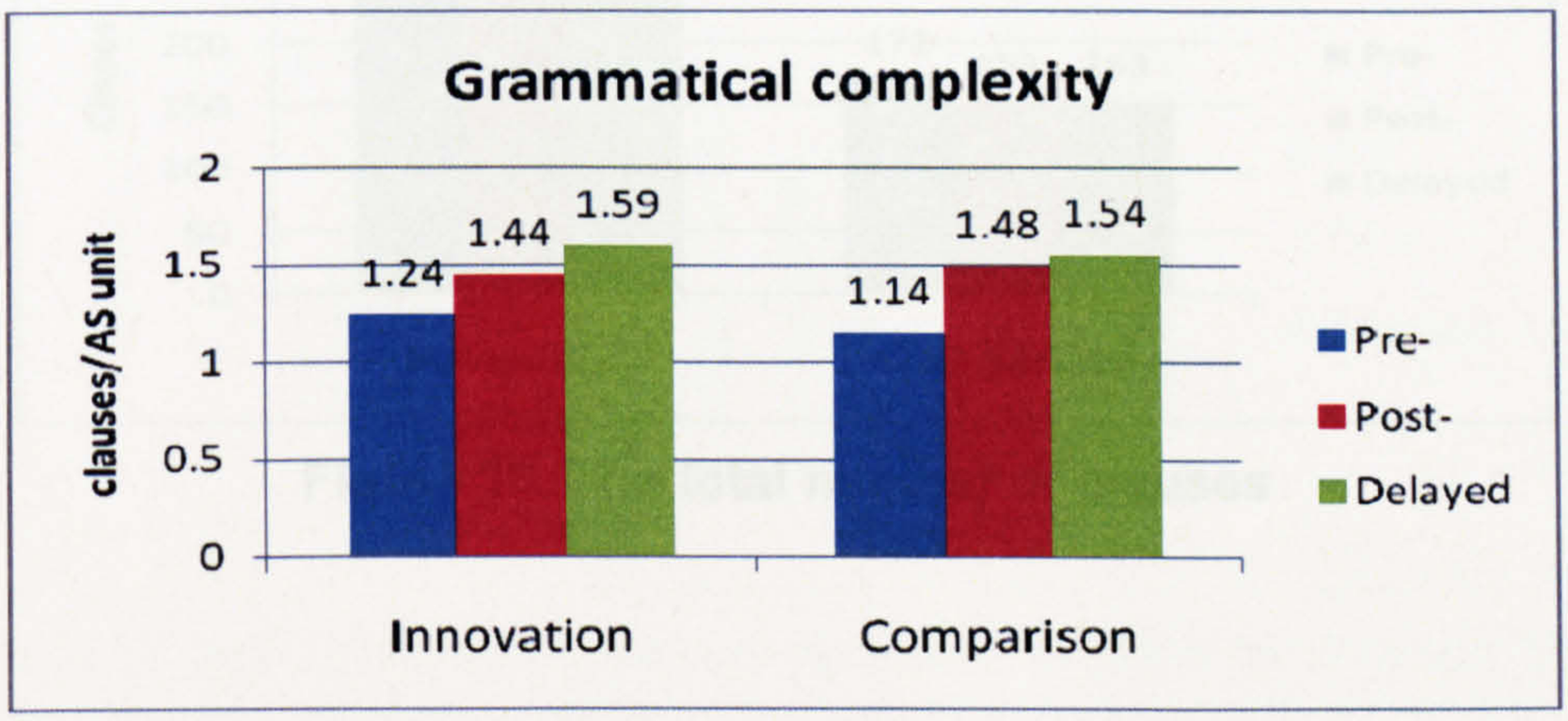


Figure 15. Grammatical complexity

Figure 15 suggests that both classes made steady improvement in grammatical complexity. However, did they genuinely improve? The total number of clauses was supposed to increase while the grammatical complexity improved. If the total number

of clauses is examined, as Figure 16 below shows, the number for the innovation class increased from 261 to 281, and to 310 at delayed post-test, while it dropped from 172 to 153 in the comparison class. This indicates, on the whole, that grammatical complexity in the comparison class did not genuinely increase as much as the figure showed. The reason is that its total number of AS units decreased significantly from 150 to 103 units at post-test, as shown in Figure 17; it is this factor which resulted in the significant increase of grammatically complexity at post-test, without improvement in the total number of clauses. Thus, the progress made in the comparison class, on the whole, was judged as relatively limited due to a decrease in the total number of clauses. By contrast, the innovation class showed a genuine improvement, where the total number of clauses increased. This analysing tool has a potential problem in comparing across different phases to determine the effectiveness. This issue will be further discussed in Chapter 6.

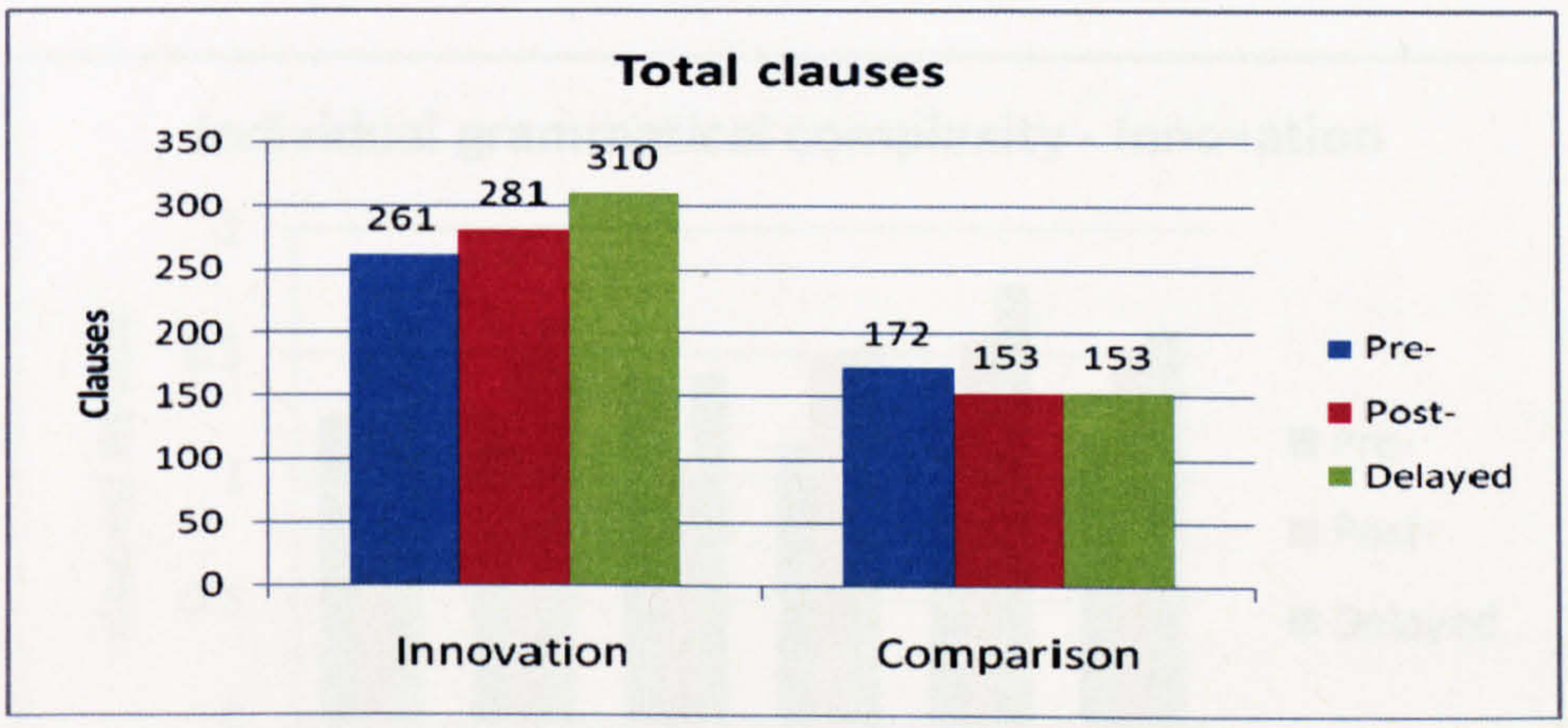


Figure 16. The total number of clauses

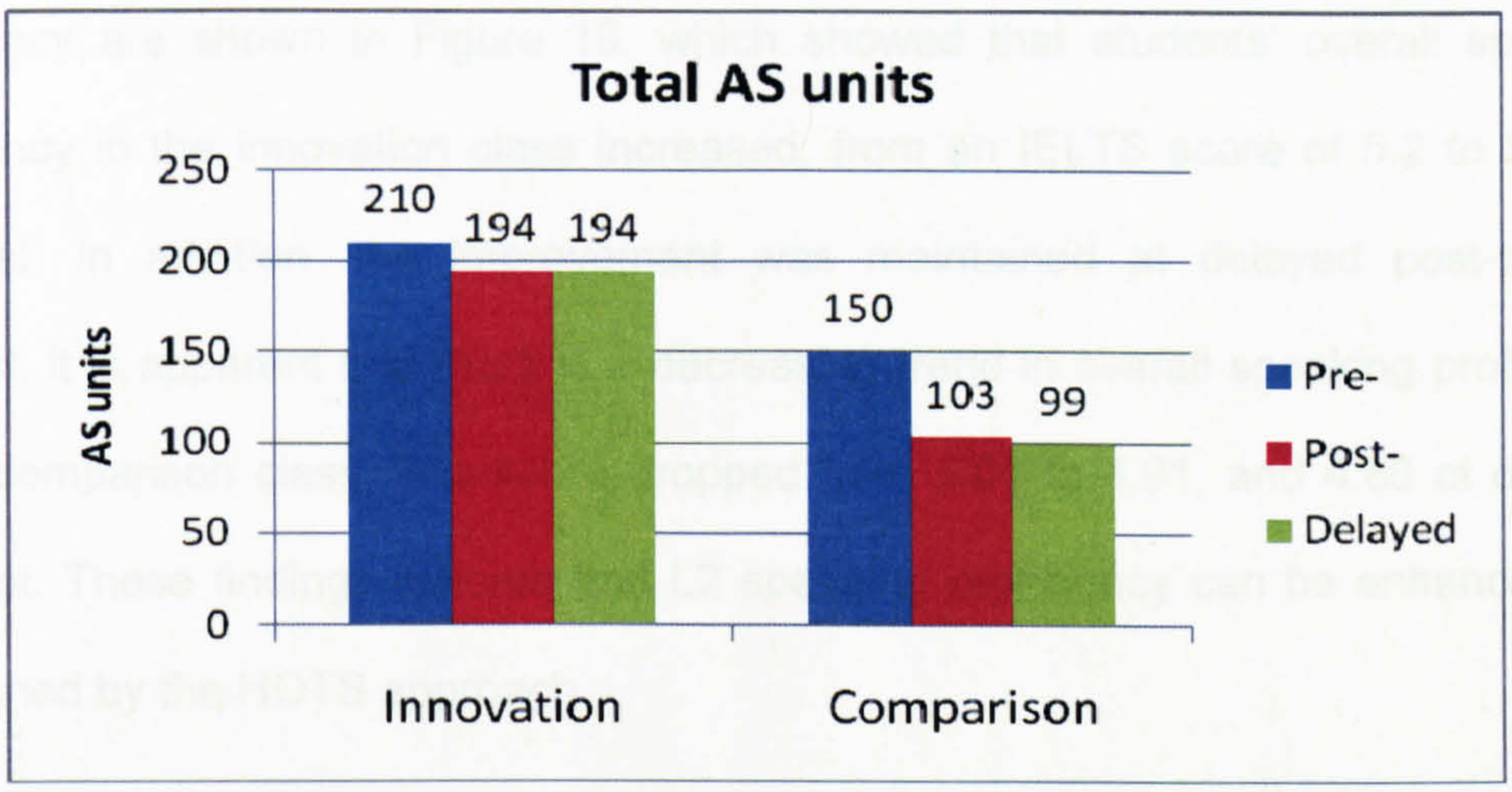


Figure 17. The total number of AS units

The improvement made in grammatical complexity in the innovation class was further justified by exploring individual grammatical complexity. Figure 18 reveals that five students progressed at post-test, and they all maintained the improvement at delayed post-test. This indicates a generic improvement made in grammatical complexity.

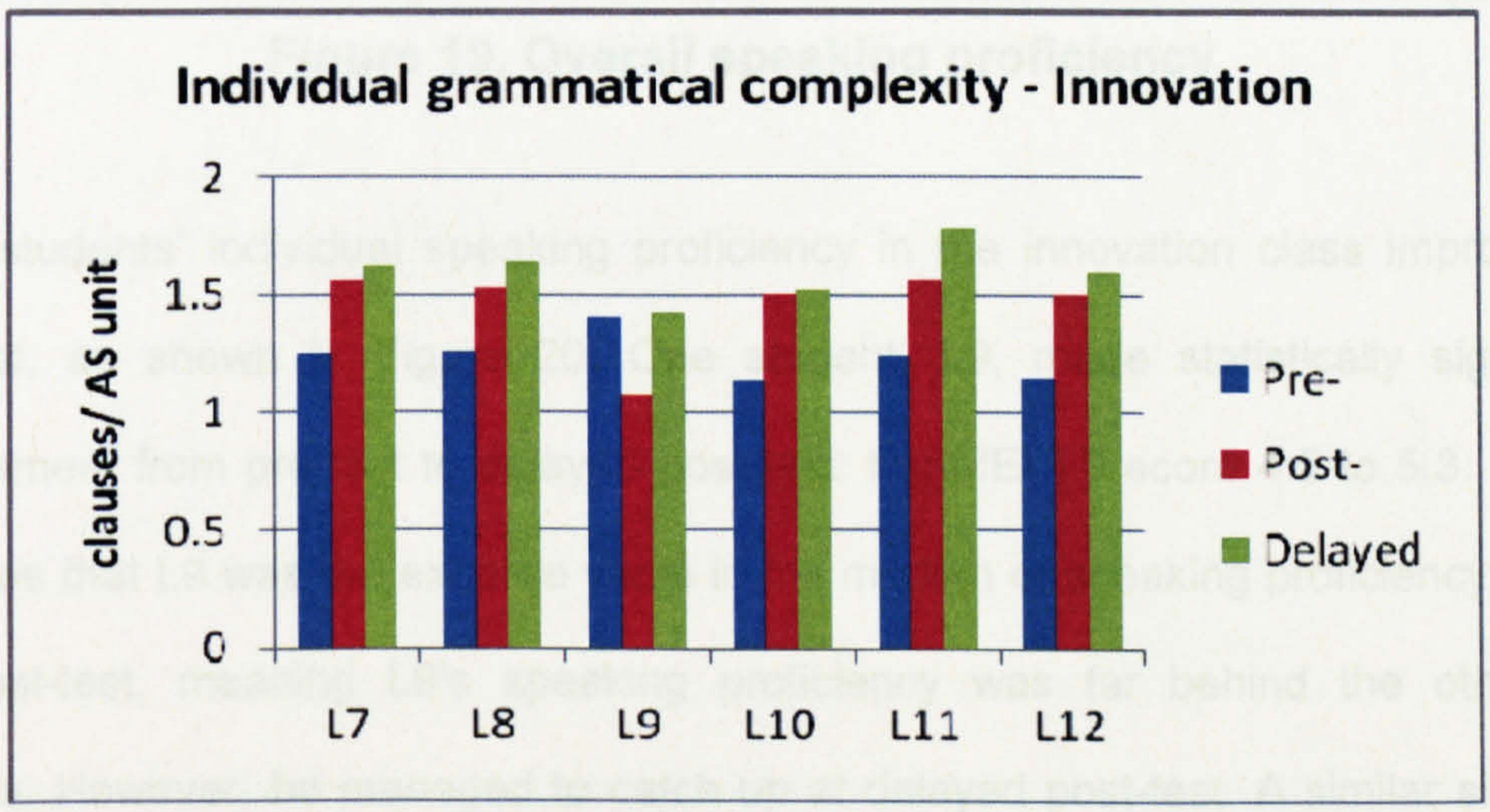


Figure 18. Individual grammatical complexity in the innovation class

Overall speaking proficiency

Did the improvement of the linguistic competence, including the length of utterance and grammatical complexity represent a good command of spoken language? This section explores this aspect. The results obtained from the analysis of speaking

proficiency are shown in Figure 19, which showed that students' overall speaking proficiency in the innovation class increased, from an IELTS score of 5.2 to 5.45 at post-test. In addition, the improvement was maintained at delayed post-test. In contrast, it is apparent that there is a decreasing trend in overall speaking proficiency in the comparison class. The score dropped from 5.01 to 4.91, and 4.88 at delayed post-test. These findings indicate that L2 speaking proficiency can be enhanced and maintained by the HOTS approach.

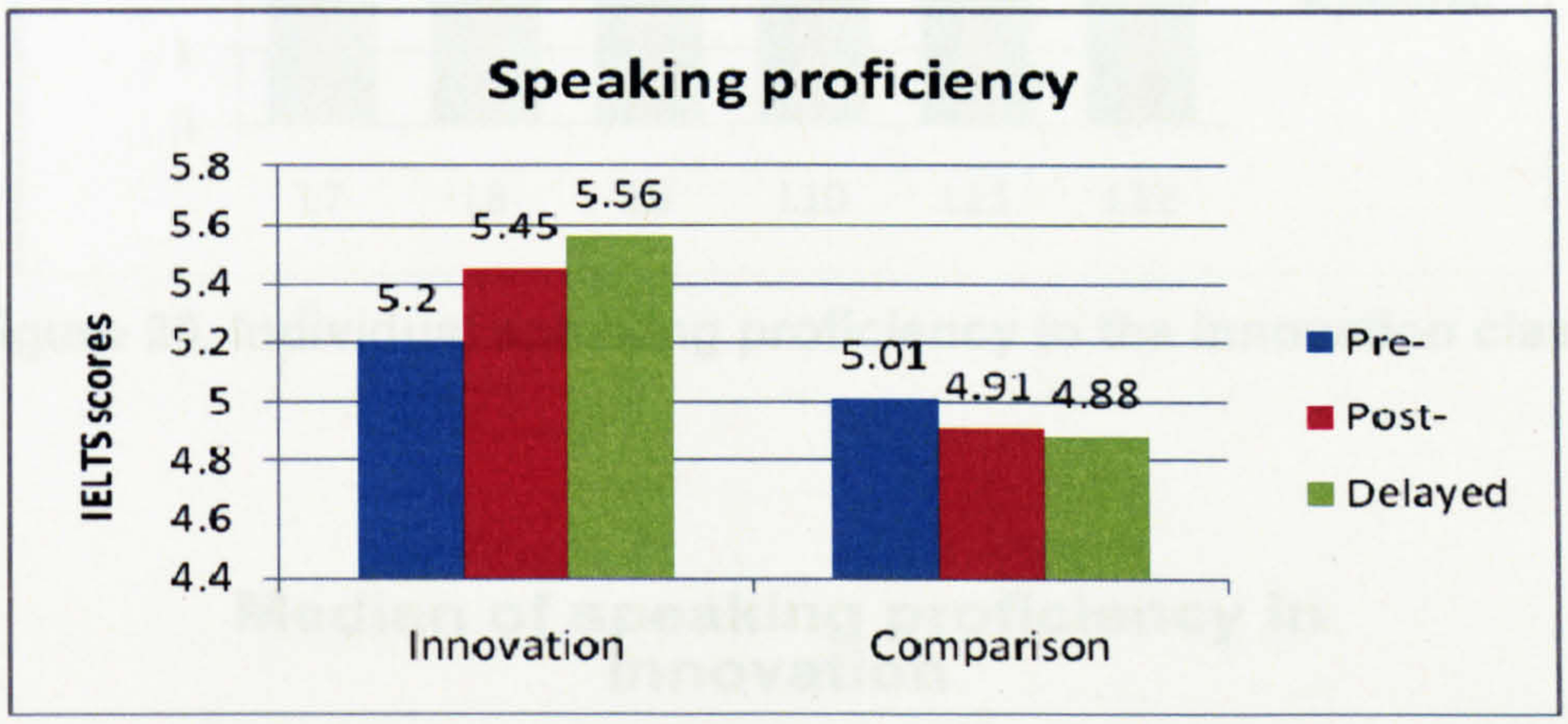


Figure 19. Overall speaking proficiency

All six students' individual speaking proficiency in the innovation class improved at post-test, as shown in Figure 20. One student, L9, made statistically significant improvement from pre-test to delayed post-test: from IELTS score 4.5 to 5.3. Figure 21 shows that L9 was the extreme value in the median of speaking proficiency at pre- and post-test, meaning L9's speaking proficiency was far behind the other five students. However, he managed to catch up at delayed post-test. A similar situtation happened with L12: although L12's improvement was not statistically significant, his IELTS score moved from 5.1 to 5.6, a considerable change. Interestingly, these two students were in different groups and were the ones with lower speaking proficiency in the group. Their improvement was even greater than the high speaking proficient students. This suggests that low speaking proficient students are able to improve their

speaking ability through the practice of thinking tasks with high speaking proficient students.

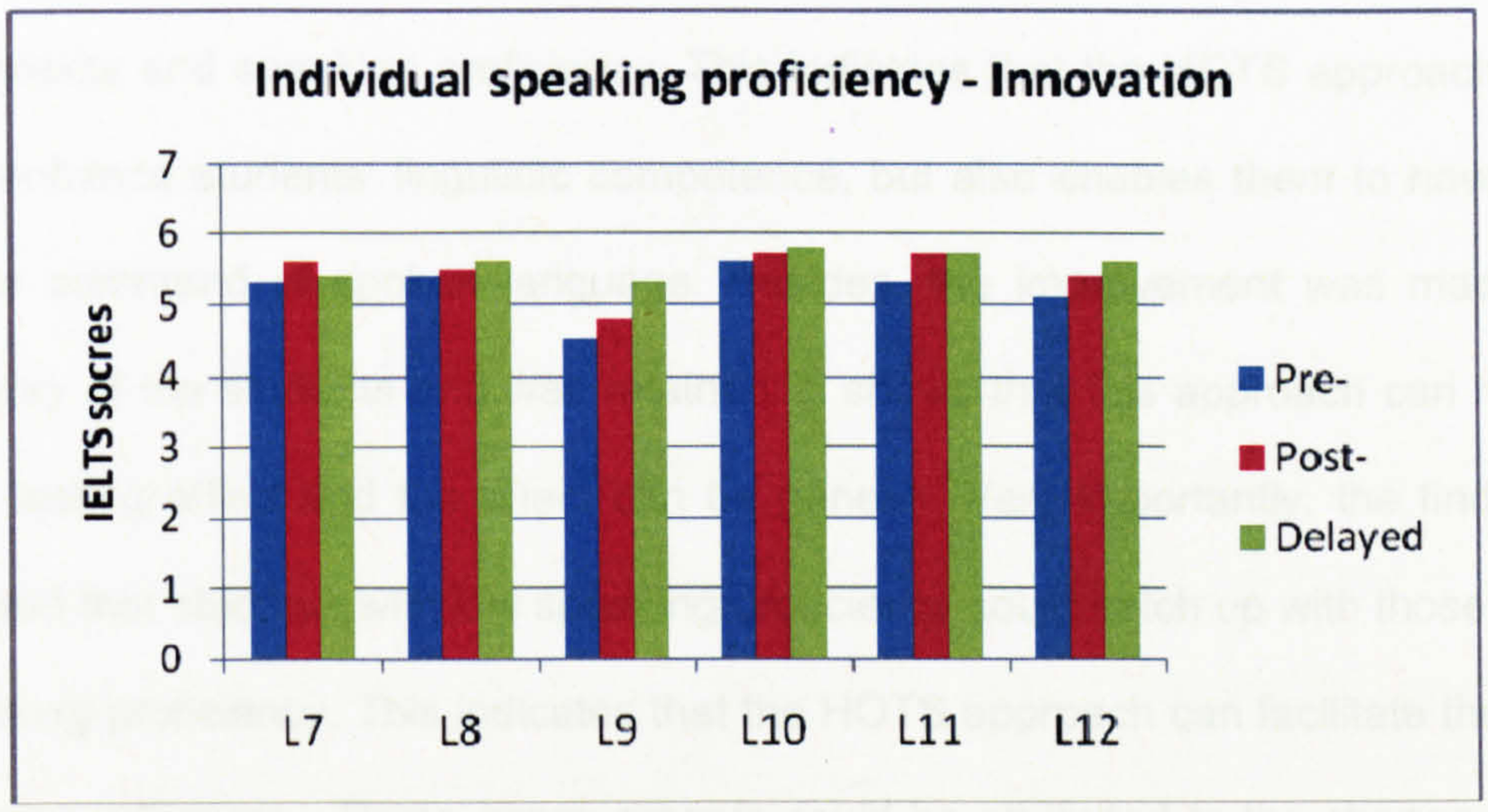


Figure 20. Individual speaking proficiency in the innovation class

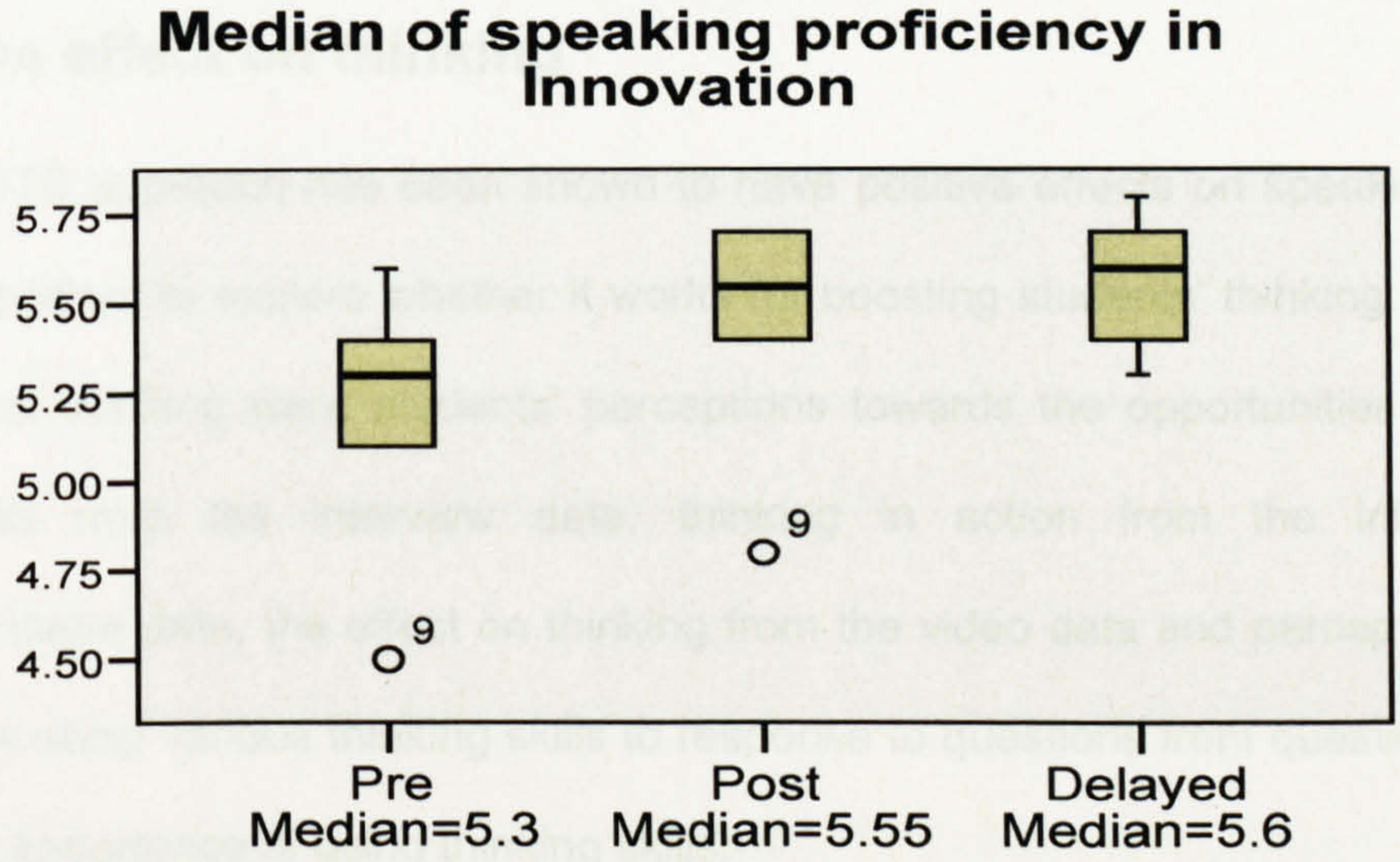


Figure 21. Median of speaking proficiency in the innovation class

4.2.4 Summary

The effects on speaking were noted from two perspectives: participants' perceptions and students' utterance recorded in the video data. The results indicate that students in the innovation class were aware of their improvement and the L2 learning process in terms of speaking opportunitites provided and their perceptions reflected the evidence

provided by the video data, such as the longer utterance produced and speaking techniques applied. In addition, the video data showed that the innovation class outperformed the comparison class in terms of the length of utterance, grammatical complexity and speaking proficiency. This indicates that the HOTS approach not only can enhance students' linguistic competence, but also enables them to have a much better command of spoken language. Besides, the improvement was made by the majority of the students and was retained. It shows that this approach can result in a long lasting effect and the effect can be generic. Very importantly, the findings also showed that students with low speaking proficiency could catch up with those with high speaking proficiency. This indicates that the HOTS approach can facilitate the learning of lower achievers. These improvements could be attributed to the change of lerning culture in the innovation class, which will be elaborated in detail in Section 4.4.

4.3 The effect on thinking

The HOTS approach has been shown to have positive effects on speaking, and it is now important to explore whether it works for boosting students' thinking. The related results of thinking were students' perceptions towards the opportunities for thinking observed from the interview data, thinking in action from the interview and questionnaire data, the effect on thinking from the video data and perceptions of their ability of using various thinking skills to response to questions from questionnaire data and the importance of using thinking skills.

4.3.1 Perceptions of the opportunities for thinking

Interviewees recognized that the innovation was cognitively demanding and four themes were found exemplifying why this innovation was perceived to make students think more frequently: the question itself, language use, listening to what others said, and extending the topic. In addition to interview data, questionnaire data also provided related findings.

Nine of the sixteen interviewees (60%) considered the thinking tasks were *intellectually challenging*. It was claimed that with this innovation they were required to think harder when answering questions and interacting with others, as stated below:

L11: 'It was cognitively demanding the first time we played *Odd One Out*.'

Interview data showed that students recognized that higher-order questions offered opportunities for thinking and they also provided reasons for this. Unlike lower-order questions, the nature of the higher-order question itself required students to think more profoundly about what it was asking. Participants claimed they devoted more effort to comprehend what the question meant, as L13 stated below.

L13: 'In the beginning of the innovation I did not understand the question and it took me a long time to think.'

This finding shows that in order to activate students' higher-order thinking skills (HOTS), they first of all needed to understand the question. Through the practice, students were able to understand the question better and thus a greater opportunity for them to think and elaborate their views.

Language used for answering the questions also required students to think about, e.g. words and sentence patterns. Learners needed to think critically and use the language creatively and logically based on the unique demand of expressing their own thoughts, as reported by L9:

L9: 'I normally think over words and then find a way to explain words I am not familiar with. Thus people understand what I mean.'

During group discussion, interaction occurred when students listened to their interlocutors. As the interviewees claimed, in order to interact with others it was

necessary for them to listen and to think about what others had said. L10 reported:

> L10: 'Students who want to learn more get more chances of learning to think... I need to think over what the others talked about.'

Extending the topic was considered an important issue in the thinking tasks. Students could control the topic; this led to more chances being created for students to think and to talk. The interviewees stated that they needed to think about different aspects relating to a topic so they could explore more of the implications, as L7 stated:

> L7: 'When a question is posed, I can explore many perspectives of that topic.'

These findings clearly indicate that during this innovation students experienced different aspects that required them to think. Interestingly, they reported that more effort was required mainly for language use and interaction, rather than figuring out the answers in their native language, Chinese. Initially students felt that the tasks were challenging, but it seemed that after practice the majority of students were able to adapt to using higher-order thinking skills (HOTS).

4.3.2 Thinking in action

This section unveils the participants' perceptions of the effect on their thinking in terms of thinking speed, socio-cognitive conflict, and the impact on their concept shaping. Alongside these effects as perceived by the interviewees, further examples extracted from the video data are provided to show how students actually engaged with HOTS in this innovation. In addition, how their concepts, e.g. values, were shaped by the tasks will also be exemplified.

It is surprising to find that four of the sixteen interviewees perceived their thinking speed to have accelerated. They claimed that at the beginning of the innovation it took them much longer to think and to express their ideas, as illustrated by the comments from L7 and L8 below:

> L7: 'It is obvious that in the beginning it took me two or three minutes to express an idea. Now it only takes two seconds.'
>
> L8: 'At the beginning I thought in Chinese, translated Chinese answers into English and then elaborated. Now I think in Chinese and elaborate immediately in English.'

Although the number of students reporting it was less than half, this finding indicates that with the HOTS approach students can be trained to process information and respond faster.

It is also interesting to find that six of the sixteen interviewees felt they were actively engaging with the use of higher-order thinking when *socio-cognitive conflict* occurred, or, as one student called it, a 'spark'. They argued with and persuaded each other in order to reach an agreement, as L8 claimed:

> L8: "The 'sparks' occurred when discussing... It is brain storming...We followed the instruction to reach an agreement in which we tried to argue with and persuade each other."

Evidence extracted from video data is provided in Excerpt 3 below as an example of this analogy of *socio-cognitive conflict*. The occurrence of socio-cognitive conflict was due to the disagreement among group members. One male student, L9, considered appearance an important feature while choosing an ideal mate. However, two female students, L7 and L8, strongly disagreed with him. As a result, there was an argument. L7 and L8 both provided reasons to support their argument (see lines 10 to 18 below)

in order to persuade L9 that appearance should not be an important feature when choosing an ideal mate.

Excerpt 3 – 'Socio-cognitive conflict'

Line	
1	L9:...appearance is quite important...maybe she look nice, pretty, cute...
2	...
3	L9: it's my favourite
4	L9: and no one try the...
5	L8: appearance, I think appearance is the most not important
6	...
7	L7: I think ::appearance is not very important
8	L9: maybe she is not very beautiful, not very pretty, but I just like something look very good
9	...
10	L7: if you deal with the people every day, you look the same
11	L8: ya
12	L7: you will don't care about what he look
13	L8: and everyone will get old and old
14	L7: ya, old
15	L8: beauty is not long time, just 10 years, or er...20 years,
16	L7: maybe your wife is a pretty and hot girl, he give you a baby, he will be fat...
17	L8: ya,
18	L7: you will love her more... because he loves you...
19	L9: my beautiful's meaning is not look nice, and someone movie star, he is not quite
20	beautiful... my meaning is I look him, my heart tell me, oh, is he, you know...

The questionnaire (Section 4, No. 2; No. 4) provided further numerical data and showed that 90% of the students believed they could express their ideas more logically, providing reasons and evidence; 72.5% of the participants considered their thinking skills had developed which meant that their HOTS had been enhanced. The findings indicate that the majority of the students perceived they had benefited from this innovation in terms of thinking performance. The occurrence of socio-cognitive conflict shows that students used the target language to communicate and argue which indicated that students' argument-based reasoning skills in English language were developed.

One result to emerge from the interview and video data was that concept shaping occurred in the learning process. In the interview data, the participants claimed that they noticed that certain words were interpreted in different ways among group members, and this resulted in a difficulty to reach an agreement during group

discussion. Thus, there was an impact on the concept of word interpretation, as stated by L11.

> L11: 'We even talked about philosophy because of that topic, since we had different concepts of the words mentioned. The problem we had could not even be solved by discussing in Chinese.'

Evidence from the video data is provided in Excerpt 4 below to exemplify this analogy of impact on word interpretation. In this discussion, L11 thought being *loyal* meant that one needed to tell friends every secret. Yet, L10 perceived being *loyal* as keeping friends' secrets and not telling others. Since they had different interpretation, L11 became confused and questioned about her own interpretation by inquiring the real meaning of the word. This shows that there was an impact on word interpretation.

Excerpt 4 - 'Concept shaping - word interpretation'

Line	
1	L12: ya, I think loyal is important...
2	L11: but, I think it's a little...I think we don't need to be loyal all the time, because everyone
3	have his own secret... sometimes he didn't need to talk to the other...
4	...
5	L10: if you tell your friend this secret, you want him or she to keep your secret...
6	L12: if the... friend is your good friend, he should be
7	L10: maybe you have many secrets, ...one of this secret you told your friend, and you don't
8	want him to say it to other...
9	L11: yes, I know, er...*what's the real meaning about loyal*...

The other impact of concept shaping was on *values*. Interviewees claimed, they discovered their values, for example, the value of competition or appearance, varied among the group members, as reported by L12 and L11:

> L12: 'I talked about that the competition could be positive competition, but they did not think so...
>
> L11: We believed that positive competition was a kind of support between friends...This is not the problem at language level anymore.'

Excerpt 5 below displays a classic example of *values* being shaped. Initially, L12 considered appearance important when choosing an ideal mate. However, L10 and L11 disagreed with L12's opinion and argument ensued. Finally, L12 was persuaded and agreed that appearance should not be seen as an important feature. The process of L12's value of appearance being shaped can be observed: from the standpoint in line 1 ('I think appearance is important because...handsome is important'), to a slightly different value in line 9 ('of course no, I think the appearance just...er... not too terrible'), and finally reaching a very different stage in line 24 ('so appearance is out'). It is through evaluating and arguing that the value was shaped.

Co-constructed reasoning, a feature of collaborative learning, occurred: one in line 3, 7 and 8, and the other in line 17 and 18. The evidence shows that high cognitive thinking can be generated collaboratively; this is why the present study did not look at individual's cognitive performances in a social context, as discussed in Section 3.8.1.

Excerpt 5 - 'Concept shaping - value'

Line	
1	L12: I think appearance is important because...handsome is important
2	...
3	L11: but I think appearance is not important...
4	L10: ya
5	L12: why?
6	...
7	L10: if he just have appearance and the other thing is bad... it's awful, then you still choose
8	him to be with you
9	L12: of course no, I think the appearance just...er... not too terrible...
10	L10: you mean he should let you looking comfortable...
11	L12: yes, yes...
12	L10: ok...
13	L12: not too handsome, or too terrible...that's ok...
14	L11: but appearance can change easy, the other character maybe is hard to change...
15	L10: the appearance is not eternal...
16	L12: that's right, so...
17	L10: maybe he have a car accident and his face is ...
18	L12: so his face is...cry
19	L10: ya...
20	L11: so you don't...
21	L10: you don't like him anymore...
22	L12: yes, I understand...
23	...
24	L12: so appearance is out...

The findings illustrate that students were aware that concept shaping occurred in the learning process. It indicates that knowledge can be co-constructed through collaborative interaction. In addition, it also shows that transformative learning occurred in the learning process.

4.3.3 The effect on thinking

The majority of students perceived that their thinking skills were developed. This section reveals whether the effect on thinking, using video data, also shows a positive result. As a reminder, cognitive performance was investigated by examining the *difficulty* level of HOTS: as determined by the frequency of the occurrence of HOTS and the function applied. Relevant examples extracted from the video data are also provided to exemplify the results.

The frequency of HOTS

The difficulty level was first determined by the frequency of HOTS. Figure 22 shows that the occurrence of lower-order thinking increased at post-test in the innovation class, while occurrences of higher-order thinking dropped significantly from 73 to 57. This finding could be seen as discouraging, but the significant decrease in the use of HOTS does not necessarily mean students' cognitive performance declined. Rather, students spent more time trying to use their comprehension thinking skills (lower-order thinking skill) to understand which interpretation of the word was correct and what the task required. Evidence is shown in the results of the detailed use of the six thinking levels in Figure 23. It shows a significant increase in the number of cases of comprehension thinking, from 32 at pre-test to 58 at post-test. However, the occurrence of higher-order thinking increased again to 83 at delayed post-test. Without any speaking practice after the innovation finished, students were able to increase the use of HOTS from 57 to 83 occurrences at delayed post-test. It infers that this ability was enhanced during the innovation, but the enhancement was demonstrated at

delayed post-test rather than at post-test due to the problem of word interpretation. As a result, the figure did not really show the maintenance.

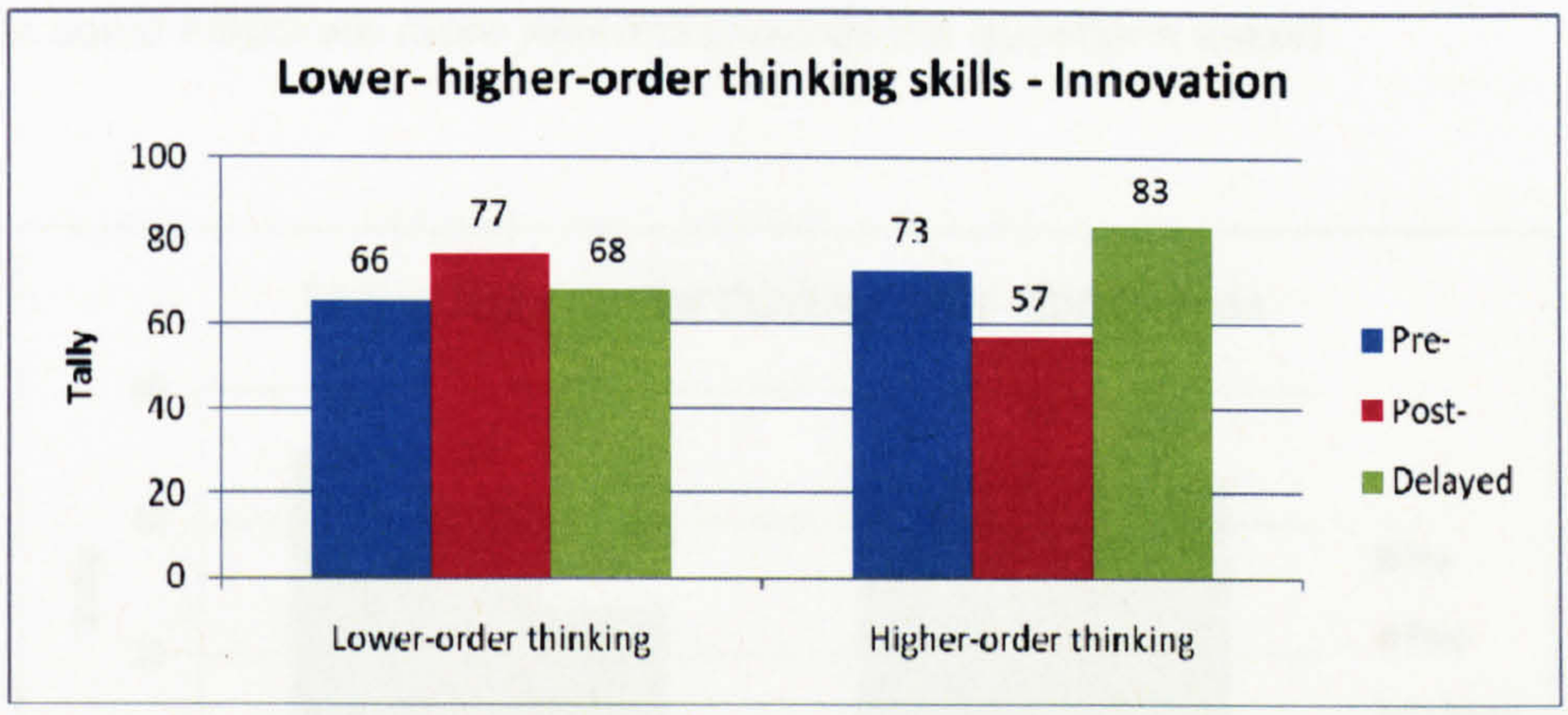


Figure 22. Lower-order and higher-order thinking skills used by the innovation class

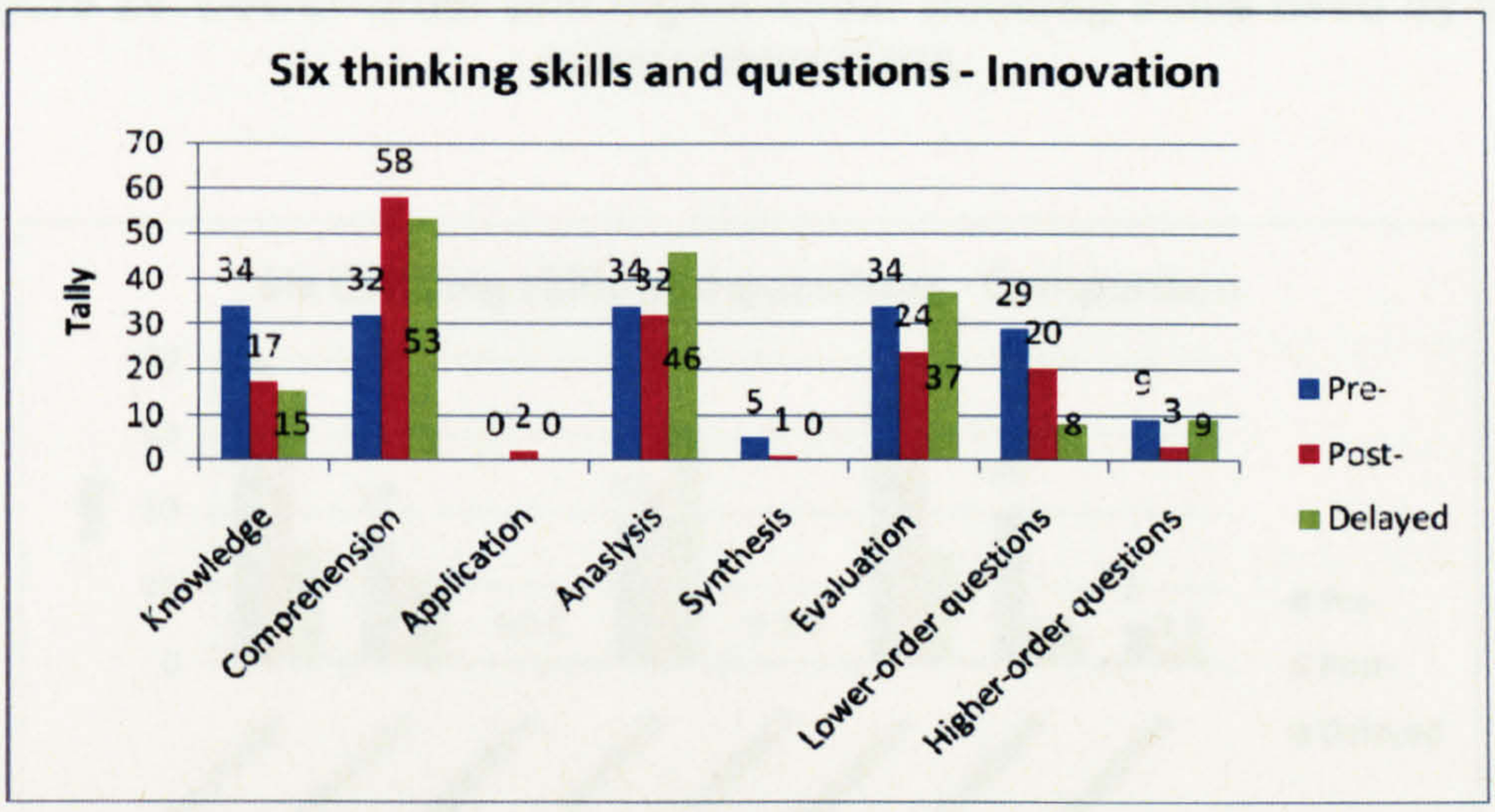


Figure 23. Six thinking skills used by the innovation class

Figure 24 presents data relating to the thinking skills used by the comparison class. A trend can be seen in the decrease in the occurrence of lower-order thinking, from 51 to 27, while the cases of higher-order thinking remained approximately the same. This indicates very little improvement in terms of the occurrence of HOTS was made.

With regard to lower-order thinking skills, the use of *knowledge* and *comprehension* thinking decreased, as shown in Figure 25. This further indicates that the comparison

class used less knowledge and comprehension thinking to understand things like what others had said or what was going on during the discussion. However, among HOTS, the use of analysis increased at post-test and delayed post-test which indicated that students could elaborate more reasons towards the questions asked.

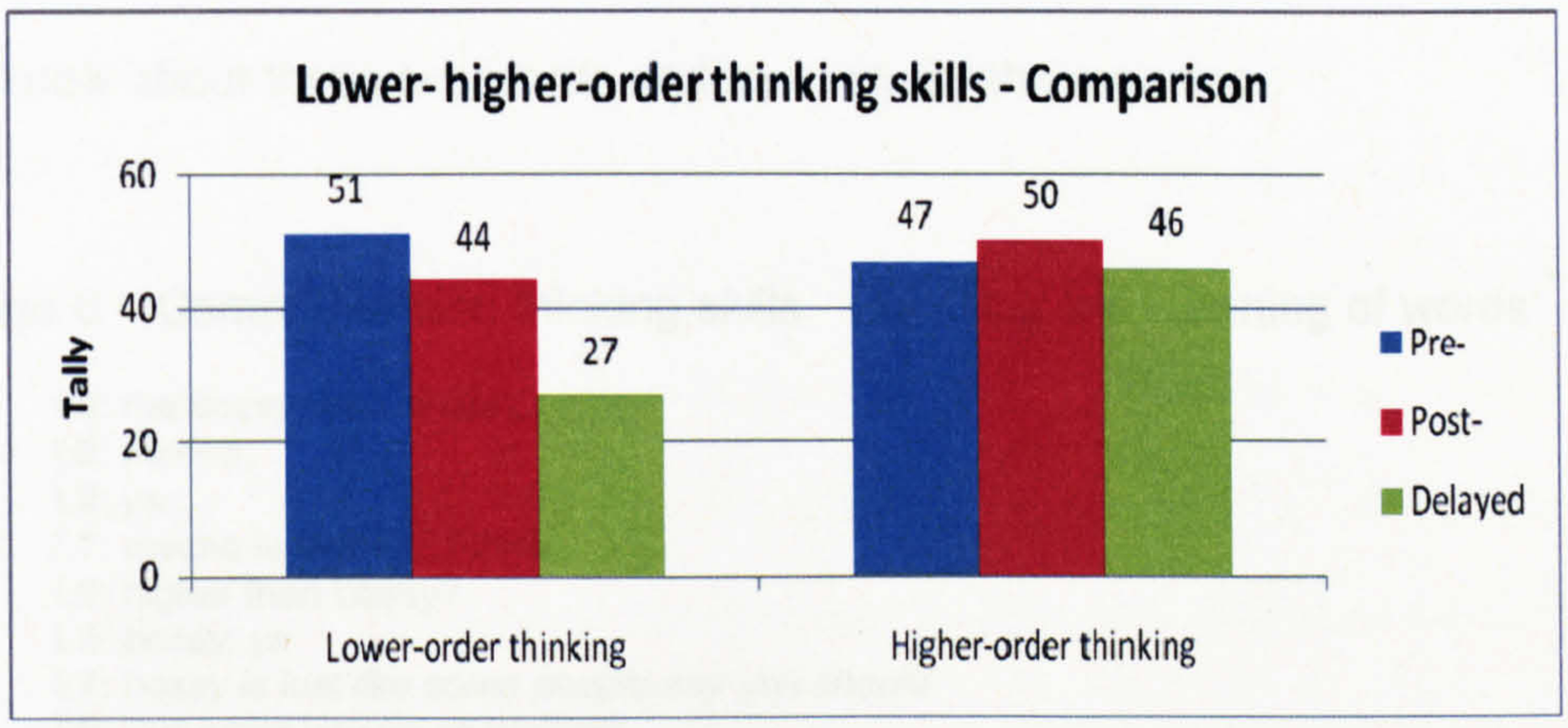


Figure 24. Lower-order and higher-order thinking skills used by the comparison class

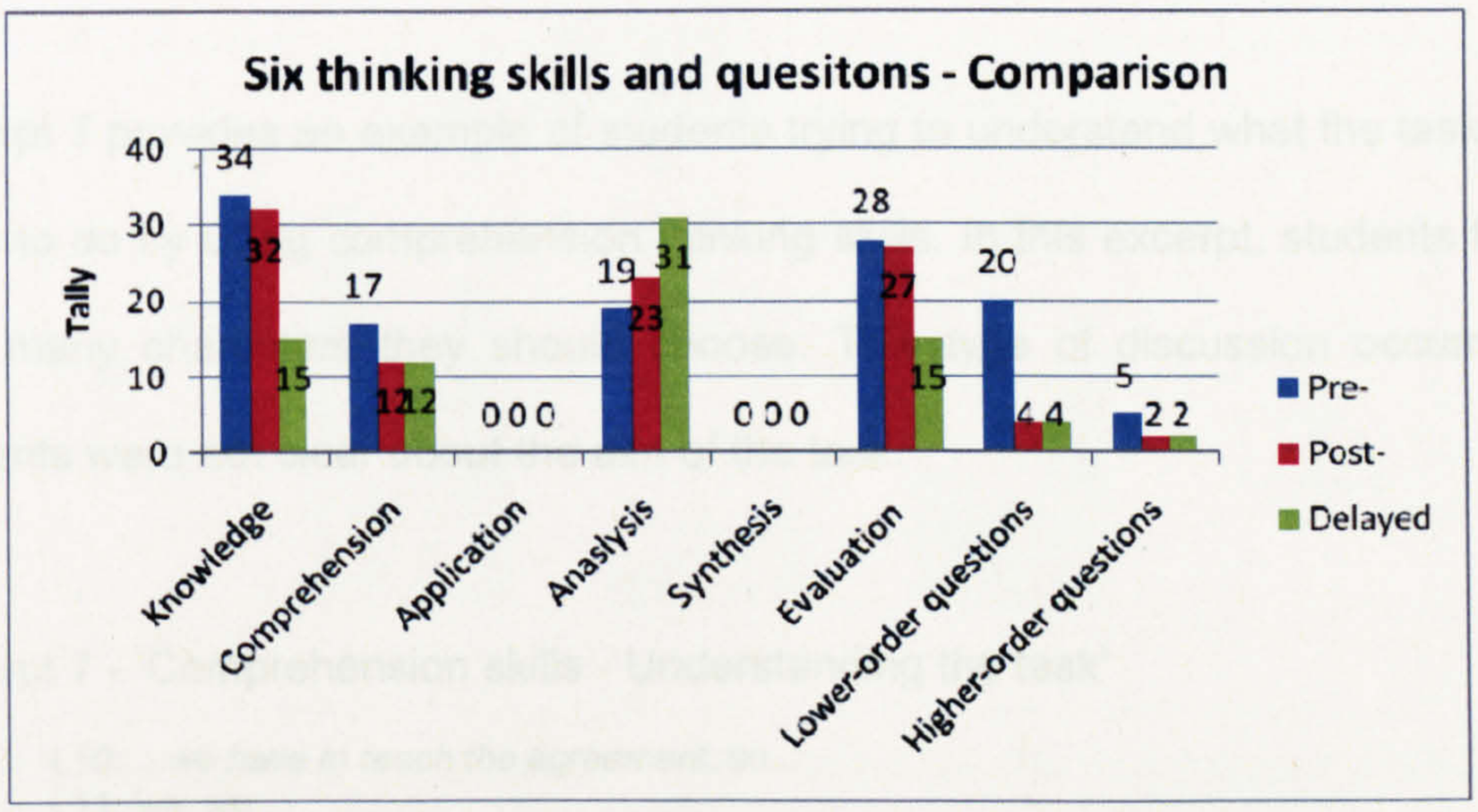


Figure 25. Six thinking skills used by the comparison class

A comaprison of the frequency of HOTS showed that the innovation class imporved more than the comparison class. Due to the problem of word interpretation, the innovation class was not able to show a maintenance of the frequency of HOTS.

The following excerpts, extracted from video data, provide examples of the increased use of comprehension thinking at post-test in the innovation class. As previously mentioned, comprehension thinking is used when one is trying to understand what is going on, to explain or to summarize. Excerpt 6 exemplifies students trying to understand the meaning of two words: bossy and macho. Students elaborated what they know about these two words and try to clarify the meaning .

Excerpt 6 - 'Comprehension thinking skills - Clarifying the meaning of words'

Line	
1	L9: *the bossy is quite like...*
2	L8: *macho,*
3	L9: *ya...*
4	L7: macho is higher... i think
5	L9: higher than bossy?
6	L8: *bossy, ya*
7	L7: *bossy is just like some people say you should*
8	L8: *just order you do something*
9	L7: *you should do something...*
10	L9: *ok control you*
11	L8: *but macho is you can't say anyting...*
12	L7: *yes...control...*

Excerpt 7 provides an example of students trying to understand what the task required them to do by using comprehension thinking skills. In this excerpt, students figure out how many characters they should choose. This type of discussion occurred when students were not clear about the aim of the task.

Excerpt 7 - 'Comprehension skills - Understanding the task'

Line	
1	L10: ...*we have to reach the agreement*, so...
2	L11: we, we
3	L10: so what's the three character we all ...
4	L12: *we need to choose one*
5	L11: *choose three...*

The findings showed that the HOTS approach had a positive impact on the frequency of occurrence of HOTS. Furthermore, the increased use of comprehension thinking in the innovation class showed students used comprehension thinking to clarify problems encountered in the discussion and to understand the task.

The function of HOTS

The function of HOTS was another means to determine the difficulty level by investigating how high cognition was used in group discussion, such as for answering questions on the handout, answering inquiries or commenting. In Figure 26, the occurrence of HOTS for interaction used in the innovation class, e.g. enquiry and comment, is shown, totalling 24, 28 and 45 through the three stages. The occurrence at delayed post test was approximately twice as many as pre-test. The figures also show this was sustained. The frequent interaction with HOTS indicates that students' ability to manage a larger amount of information was enhanced and this shows the difficulty level of HOTS increased. Being able to manage a larger amount of information within a certain amount of time also implies that students' thinking speed was accelerated. In addition, the findings show peer interaction with high cognition developed.

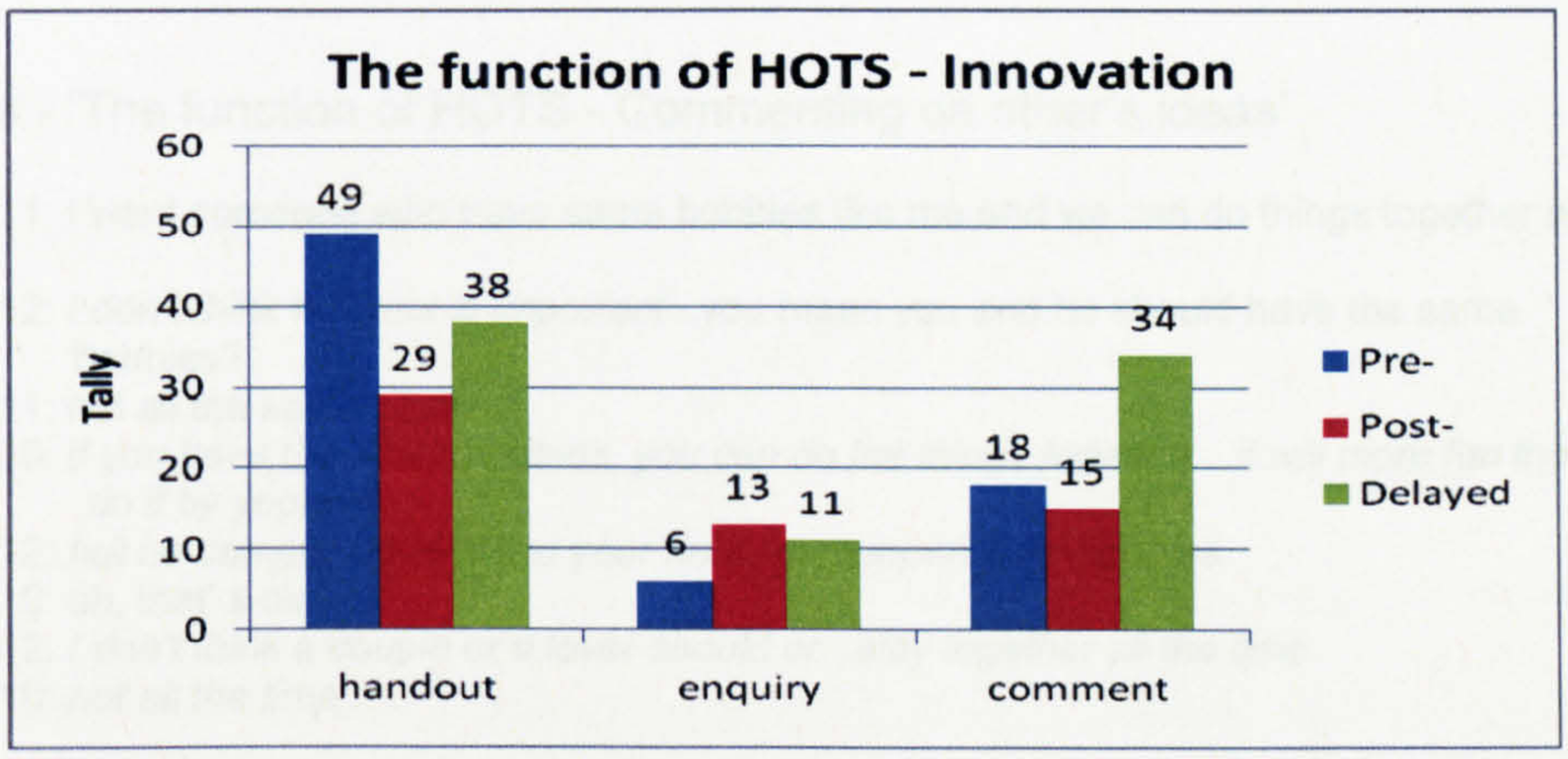


Figure 26. The function of HOTS used by the innovation class

Figure 27 presents the function of HOTS used by the comparison class. It shows that the HOTS ideas expressed were mainly for answering questions listed on the handout, while ideas expressed for answering enquiries or for commenting on others' opinions were relatively low. The findings infer that HOTS were mainly used in 'one-way' expression without interacting with others.

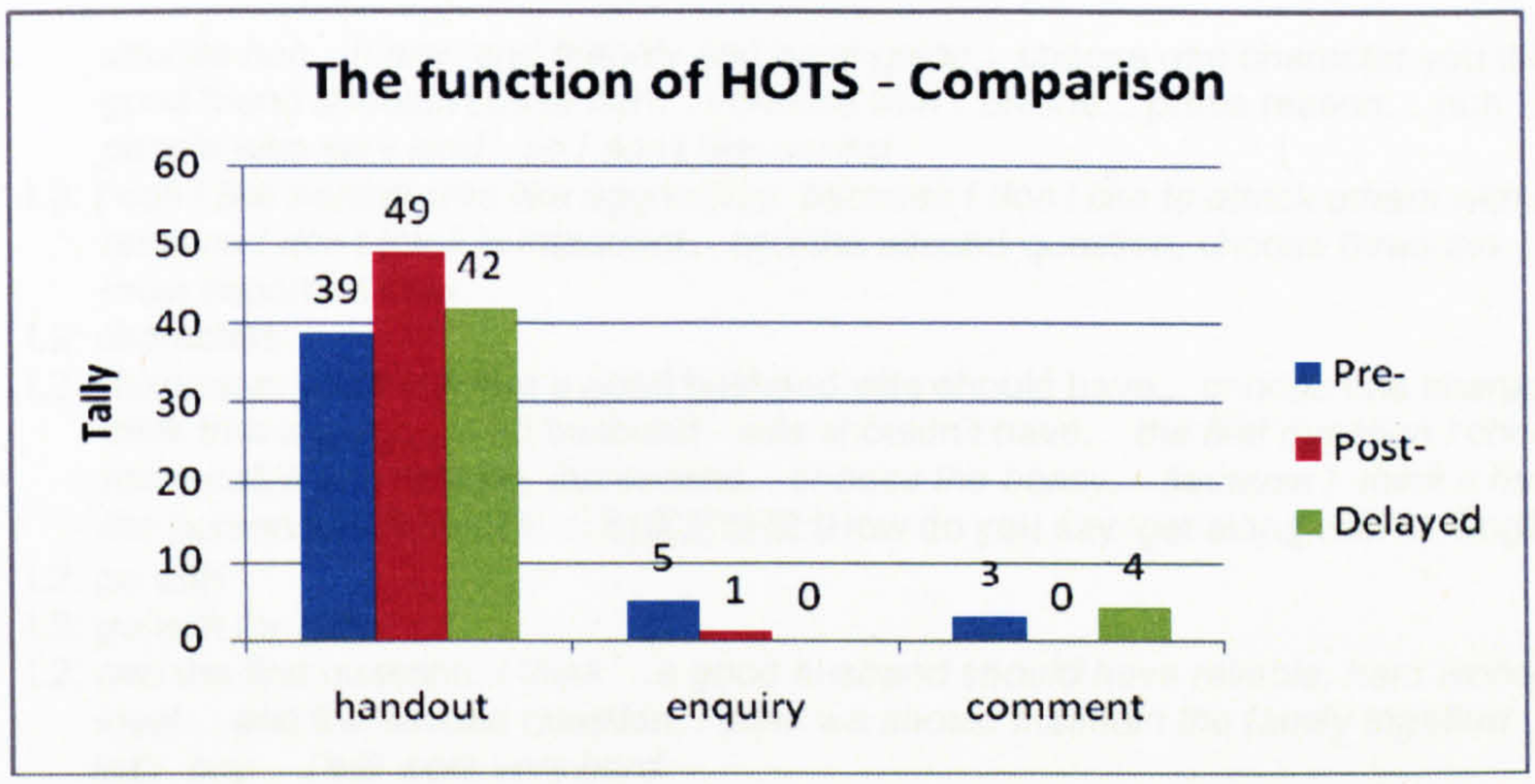


Figure 27. The function of HOTS used by the comparison class

An example of the ideas expressed for commenting or further building on others' opinions is shown in Excerpt 8, extracted from the video data collected in the innovation class. Students had different opinions regarding the importance of *hobbies* when choosing an ideal mate. They argued and made comments, as shown in the words in italics, and the talk was co-constructed.

Excerpt 8 - 'The function of HOTS - Commenting on other's ideas'

Line	
1	L11: I want someone who have same hobbies like me and we can do things together and…
2	…
3	L12: *I don't think hobbies is important*...you mean you and he should have the same
4	hobbies?
5	L11: not all the same…but
6	L10: *if you have the same hobbies, you can do the things together…it will more fun than you*
7	*do it by yourself…*
8	L12: *but he can just understand your hobbies, support your hobbies*
9	L10: oh, that' s ok…
10	L12: *I don't think a couple or a lover should er...stay together all the time*
11	L10: *not all the time…*

A representative example of the ideas expressed for answering questions on the handout in the comparison class is shown in Excerpt 9. Students read and answered questions individually, as shown in the words in italics. Without requests for, comments on, or extensions on others' ideas, the talk was *monologue*.

Excerpt 9 - 'The function of HOTS - Answering questions'

Line	
1	L1: choose the three most important characters you think that a good friend should have... /

choose huh…funny, and friendly and easy-going... choose one character you think that a good friend shouldn't have huh... *I choose huh…unkind*... prove reason... huh... *I like people who very kind…so I don't like unkind*

L3: *I don't like person who like aggressive because I don't like to attack others with the reasons I don't think is important*... and the second question, choose three the most important char...

L2: characters...

L3: characters you think that a good husband wife should have... choose one character you think that a good...good husband - wife shouldn't have... *the first question I choose kind, reasonable and reliable.* the second, *I choose the bossy… because I think a husband… is the person who I have to* …相處怎麼講 (How do you say 'get along with' in English)

L2: go with

L3: *go with for a life*

L2: and the first question, *I think …a good husband should have reliable, hard working and loyal*... and the second question, *I think we should maintain the family together… if he is a lazy boy… I will work very hard*

The results of the function of HOTS show that HOTS were more often used for interaction in the innovation class, while the comparison class rarely had interaction. Findings obtained from the frequency of HOTS and the function of HOTS indicate that the innovation class improved considerably in the ability to manage the *difficulty* level of HOTS when compared to the comparison class. Regarding whether this ability can be maintained, the results show that students were able to sustain the skill of using HOTS for interaction. However, the findings obtained from the analysis of six thinking levels did not show maintenance. These two results could be seen as rather contradictory. Yet, it could be that the problem of word interpretation that occurred at the post-test data collection affected the results of the frequency of HOTS. This explanation arguably makes sense of the finding, the maintenance, found in the analysis of the function of HOTS. On the other hand, it could be concluded that the accelerated thinking in the innovation class could be explained by a rise in the difficulty level of HOTS within a certain amount of time where students were able to manage a greater amount of information.

4.3.4 Perceptions on the ability and using thinking skills

This section first presents results looking at the impact of this innovation on students' perceptions of their ability to answer questions in the L2 classroom. Questionnaire

data from Section 2 were analysed by ANCOVA test and Paired-Samples T Test are used to triangulate.

Regarding the perceived ability to answer lower-order and higher-order questions, the ANCOVA test shows P = .494 and .136 respectively, as shown in Table 5 and 6. This means that there was no significant difference between the two classes with regard to students' perception of their ability to answer lower-order and higher-order questions, thus there was no evidence of impact from the innovation. In Table 7 and 8, Paired-Samples T Tests showed no significant figures with regard to the ability to answer lower-order (P = .200 and .130) and higher-order questions (P = .420 and .102) in each class. This indicates that from pre-test to post-test, neither class perceived any significant change in their ability to answer questions.

Table 5. Analysis of One-way ANCOVA on the comparison of the perceived ability to answer lower-order questions

Homogeneity of with-in regression of One-way ANCOVA

Some of Variation	SS	DF	MS	F	P
Between groups x covariant	.215	1	.215		
Error	53.981	65	.830	.259	.613

The analysis summary of One-way ANCOVA

Some of variation	SS	DF	MS	F	P
Covariant	.026	1	.026		
Between Groups	.389	1	.389	.474	**.494**
Within group	54.196	66	.821		
Total	572.284				

Table 6. Analysis of One-way ANCOVA on the comparison of the perceived ability to answer higher-order questions

Homogeneity of with-in regression of One-way ANCOVA

Some of Variation	SS	DF	MS	F	P
Between groups x covariant	.098	1	.098		
Error	65.893	66	.998	.098	.755

The analysis summary of One-way ANCOVA

Some of variation	SS	DF	MS	F	P
Covariant	.719	1	.719		
Between Groups	2.246	1	2.246	2.281	**.136**
Within group	65.991	67	.985		
Total	730.815				

Table 7. Paired-Samples T Test on the perceived ability to answer lower-order questions

	Paired Differences					t	df	Sig. (2-tailed)
	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference		Mean	Std. Deviation	Std. Error Mean
	Lower	Upper	Lower	Upper	Lower	Upper	Lower	Upper
Innovation	.24217	1.16000	.18575	-.13386	.61819	1.304	38	**.200**
Comparison	.36667	1.28824	.23520	-.11437	.84770	1.559	29	**.130**

Table 8. Paired-Samples T Test on the perceived ability to answer higher-order questions

	Paired Differences					t	df	Sig. (2-tailed)
	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference		Mean	Std. Deviation	Std. Error Mean
	Lower	Upper	Lower	Upper	Lower	Upper	Lower	Upper
Innovation	.18234	1.39536	.22344	-.26999	.63466	.816	38	**.420**
Comparison	.40860	1.34818	.24214	-.08592	.90312	1.687	30	**.102**

To further describe the phenomenon of how students evaluated their ability to answer questions, descriptive results were provided. Table 9 shows that the progress in answering higher-order questions perceived by students of the comparison class was much higher than that of students in the innovation class; the comparison class rated themselves at 3.68 at pre-test and 3.25 at post-test, while the innovation class rated 3.11 at pre-test and 2.91 at post-test (the questionnaire scale ranged from 1 = strongly agree, to 6 = strongly disagree). The comparison class considered themselves as having improved by 0.43, while the innovation group considered their progress as 0.2. Although there was no statistical significance between the two classes, it is worth pointing out that the mean score results indicate that the comparison class was more satisfied with their ability to answer higher-order questions than the innovation class. However, these results contrast with the findings relating to the concrete evidence of thinking obtained from video data (Section 4.3.3), in which the innovation class demonstrated a greater improvement in using higher-order thinking skills than the comparison class. This result could indicate that the HOTS approach enables students

to be more aware of what they cannot do and become more cautious when evaluating their achievement.

Table 9. Mean scores of the perceptions of the ability to answer higher-order questions

Innovation Pre-	N	39
	Mean	**3.11**
	Std. Deviation	.83863
Innovation Post-	N	40
	Mean	**2.91**
	Std. Deviation	1.00387
Comparison Pre-	N	32
	Mean	**3.68**
	Std. Deviation	.87448
Comparison Post-	N	31
	Mean	**3.25**
	Std. Deviation	.96736

Students themselves had evaluated the achievement made in thinking. It is now worth to explore the effect of this innovation on their perceptions of the importance of using thinking skills in L2 classrooms. Questionnaire data, Section 1, were analysed by ANCOVA test and Paired-Samples T Test.

Table 10 shows there was no impact of this innovation on the perceived importance of using lower-order thinking between the innovation and the comparison classes, in which P = .063. Yet, there was a significant difference in the importance of using higher-order thinking. Table 11 shows that P value was .026 ($p < .05$), meaning that after the innovation there was a significant difference in the importance of using higher-order thinking in L2 classrooms between the two classes. To further explore this phenomenon, a descriptive result was provided in Table 12. It shows that the mean scores of the importance of using higher-order thinking were 2.42 at pre-test and 2.43 at post-test in the innovation class, and 2.37 at pre-test and 2.04 at post-test in the comparison class. The mean scores reveal that in the innovation class students' attitudes towards the importance of using higher-order thinking remained

approximately the same after the innovation, while students in the comparison class considered it more important.

Table 10. Analysis of One-way ANCOVA on the comparison of the perceived importance of using lower-order thinking

Homogeneity of with-in regression of One-way ANCOVA

Some of Variation	SS	DF	MS	F	P
Between groups x covariant	.004	1	.004		
Error	22.239	68	.327	.012	.913

The analysis summary of One-way ANCOVA

Some of variation	SS	DF	MS	F	P
Covariant	.168	1	.168		
Between Groups	1.155	1	1.155	3.582	**.063**
Within group	22.243	69	.322		
Total	318.025				

Table 11. Analysis of One-way ANCOVA on the comparison of the perceived importance of using higher-order thinking

Homogeneity of with-in regression of One-way ANCOVA

Some of Variation	SS	DF	MS	F	P
Between groups x covariant	.291	1	.291		
Error	38.695	65	.595	.489	.487

The analysis summary of One-way ANCOVA

Some of variation	SS	DF	MS	F	P
Covariant	.142	1	.142		
Between Groups	3.043	1	3.043	5.152	**.026**
Within group	38.987	66	.591		
Total	393.321				

Table 12. Mean score of the importance of using higher-order thinking

Innovation Pre-	N	38
	Mean	**2.42**
	Std. Deviation	.69517
Innovation Post-	N	40
	Mean	**2.43**
	Std. Deviation	.63525
Comparison Pre-	N	31
	Mean	**2.37**
	Std. Deviation	.75498
Comparison Post-	N	32
	Mean	**2.04**
	Std. Deviation	.88383

The mean score results show that basically both classes were of the opinion that higher-order thinking played an important role in L2 classrooms. However, it is interesting to see that in the innovation class students did not consider the use of higher-order thinking more important; in contrast, the comparison class, considered it

much more important. This suggests that the comparison class may have perceived that they benefited from the tasks conducted at pre-test and post-test; thus, they considered the use of higher-order thinking to be very important and had a desire to use higher-order questions.

Interestingly, Table 13 and 14 show the results of Paired-Samples T Test on the perceived importance to use lower-order and higher-order thinking in the innovation and comparison classes. There were no significant figures, except for the use of lower-order thinking in the innovation class, P value of .028 ($p < .05$), which indicates that students' perceptions of using lower-order thinking was significantly different after the innovation. Descriptive results were provided to further illustrate this phenomenon in Table 15 in which the mean score at pre-test was 1.88, and 2.14 at post-test. The findings show that students still considered the use of lower-order thinking important; however, they perceived it as significantly less important after the innovation was conducted.

Table 13. Paired-Samples T Test on the perceived importance to use lower-order thinking

Class	Paired Differences					t	df	Sig. (2-tailed)
	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference		Mean	Std. Deviation	Std. Error Mean
	Lower	Upper	Lower	Upper	Lower	Upper	Lower	Upper
Innovation	-.25556	.70954	.11219	-.48248	-.02863	-2.278	39	**.028**
Comparison	.19097	.84301	.14902	-.11296	.49491	1.281	31	.210

Table 14. Paired-Samples T Test on the perceived importance to use higher-order thinking

Class	Paired Differences					t	df	Sig. (2-tailed)
	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference		Mean	Std. Deviation	Std. Error Mean
	Lower	Upper	Lower	Upper	Lower	Upper	Lower	Upper
Innovation	-.02339	.93219	.15122	-.32979	.28301	-.155	37	.878
Comparison	.34767	1.24383	.22340	-.10857	.80391	1.556	30	.130

Table 15. Mean score of the perceived importance to use lower-order thinking in the innovation class

Class		Pre-	Post-
Innovation	N	40	40
	Mean	**1.88**	**2.14**
	Std. Deviation	.42254	.53091

4.3.5 Summary

Participants perceived there were more opportunities for thinking and they were able to distinguish the factors that required them to think more. They also perceived their thinking to become quicker. The occurrence of socio-cognitive conflict indicates that students' argument-based reasoning skills in English and problem solving ability had developed. Knowledge can be restructured, and this indicated transformative learning took place in the learning process while the thinking tasks were in progress.

The evidence obtained from the video data shows that the innovation class improved considerably more than the comparison class in cognitive performance. The results showed that higher-order thinking skills were used not only for answering questions, but also for interaction. This is a very important finding. It indicates that the HOTS approach enhanced peer-interaction talk with high cognition which will be discussed as the term, high cognitive interactive talk, in Chapter 5. The results also show that students were able to maintain this use of high cognition for interaction which indicates that this approach can result in a long term effect thus adding to research in the field.

Regarding students' perceptions of their ability to answer lower-order and higher-order questions, the results of ANCOVA test show that there was no impact in this area. However, the descriptive results revealed an interesting phenomenon: the comparison class members perceived their progress in answering higher-order questions as being much higher than members of the innovation class perceived theirs to have been. However, this result obtained from the questionnaire data contrasted with the results gained from the video data. This implies that the HOTS approach can enable students

to be more aware of what they cannot do and become more cautious when evaluating their achievement. Having explored students' speaking and thinking performance, it is now important to examine students' attitudes towards and perceptions of this innovation.

4.4 Attitudes and perceptions

Section 4.2.2, 4.3.2 and 4.3.4 have looked at students' perceptions particularly on speaking and thinking performance. This section further reports interview and questionnaire results relating to whether they liked this innovation, perceptions of the effects on their performance other than speaking and thinking and an evaluation on the innovation tools conducted. Prior to reporting student attitudes, it is important to reveal the change of learning climate in the classroom, so that a better understanding towards student attitudes can be drawn.

The class started at half-past three in the afternoon and students often looked tired. In the beginning of the innovation, students showed a preference for speaking activities in drills rather than making their own sentences due to the extra effort required in formulating sentences. Moreover, they needed to adapt to the innovator and new teaching approach, which required them to break their previous learning habit: from a more passive oriented learning style to a more active one. Thus, they were under pressure, in particular, in a teacher-led setting. After a higher-order question was posed, the learning atmosphere changed: a nervous tension arose and students started to think hard. No one would volunteer to answer the question: students, as what they normally did in class, kept quiet unless being nominated. The researcher had to randomly select a student to answer the question or to comment on others' opinions. Although they did not participate proactively at this stage, they started to learn the interaction pattern and the value of their opinions being appreciated in a discussion forum: their ideas being heard, responded to, commented on and valued.

Students perceived such value vividly in a group discussion forum where they argued with each other in order to reach a consensus. However, one case observed showed that when the wait-time was not sufficient for students to elaborate on their thoughts, they could get upset due to their perception that ideas were not valued.

Students did not volunteer to talk in a teacher-led setting. Their motivation towards speaking, however, was shown clearly when the teaching forum was switched to group discussion: the classroom would be suddenly full of noise, and this seemed that they desperately wanted to talk about their thoughts or to work on the particular task of that week. Such noise did not occur in the beginning of the innovation; it appeared more markedly as the innovation progressed and can be seen as a switch to exploratory talk (Mercer 2000). This was attributed to the modelling conducted in a teacher-led setting where students learned how to interact in a discussion. The researcher noticed that at the commencement of the innovation, students' disposition and interaction skills, including language use for interaction, were underdeveloped: they mainly focused on elaborating upon their own ideas rarely commenting on each other's thoughts. Here, the coherence of communication was not clearly demonstrated. Two months later, their ability to build on others' ideas and make judgements and comments was more evidently demonstrated. With such a change of learning culture, the formal lecturer, upon seeing the students' discussion, noticed that their speaking ability had improved. At this stage, the coherence was promising and reasoning skills became more sophisticated.

In a teacher-led setting the proactive participation was not seen until the end of the innovation approached. Students were reluctant to talk despite having the answer to the question in mind, due to high levels of anxiety when speaking in front of the class. The anxiety seemed to decrease gradually and some students would start volunteering to answer questions. Such decreasing anxiety was also shown in group

discussion. Some students who rarely talked in class began to take courage and experiment with segmented utterances. This revealed an increasing willingness to take risks when speaking in the class.

Among the four types of tasks conducted, participants were highly motivated by, actively participated in and enjoyed the task 'Guess What I Say'. They came to the front of the class and talked regardless of their varying speaking levels, and the class was full of laughter. The formal lecturer observing the class was surprised to see the students' high levels of motivation. This could be attributed to the material designed.

Another change of atmosphere observed was that students seemed to think faster. The researcher would allow some time for students to think about the question and formulate their answers before nominating someone to talk. At the initial stage, the researcher still had to wait for a while after nominating students to provide the answer. But, gradually students could talk soon after being chosen. Even when students volunteered to talk, the post-question wait-time became markedly shorter than previously.

The difference in learning culture between the innovation and comparison classes was varied in some aspects. Regarding teacher behaviour, more higher-order questions and probing follow-up questions were posed to allow students to elaborate their thoughts in the innovation class. In addition, the techniques and language use of interaction in a social context was also modelled, which led to the occurrence of interaction with high cognition, while such teacher behaviour was rarely evident in the comparison class and thus the interaction was not developed in the same way. Unlike the innovation class, the atmosphere in the comparison class was more relaxed and the nervous tension rarely appeared; they had the same lecturer with the same teaching method employed in the previous semester which they were familiar with. In

addition to the different speaking tasks employed between these two classes, the questioning behaviour in a teacher-led setting resembled inquisition without genuine interaction taking place in the comparison class. Despite allowing students to elaborate on their thoughts, students did not comment or build on each others' ideas. This also happened in the group discussion forum; students mainly focused on providing individual answers rather than listening to and considering others' opinions. Such a learning culture could be attributed to a lack of modelling of the interaction skills.

4.4.1 Attitudes towards and perceptions of the HOTS approach

Students' attitudes towards the innovation conducted were mainly extracted from the interview data, although results gained from the questionnaires are also provided for triangulation where available. Overall, the participants considered this innovation to be characterised by *interest*, *enjoyment*, *pressure and fear*, *hard work, resistance and difficulties encountered* as shown in Figure 28. The first two were positive attitudes and could be seen as the reasons why they liked this innovation, while the last two were considered as negative attitudes. Hard work and pressure and fear were deemed as both positive and negative and these contrasting perspectives will be explored.

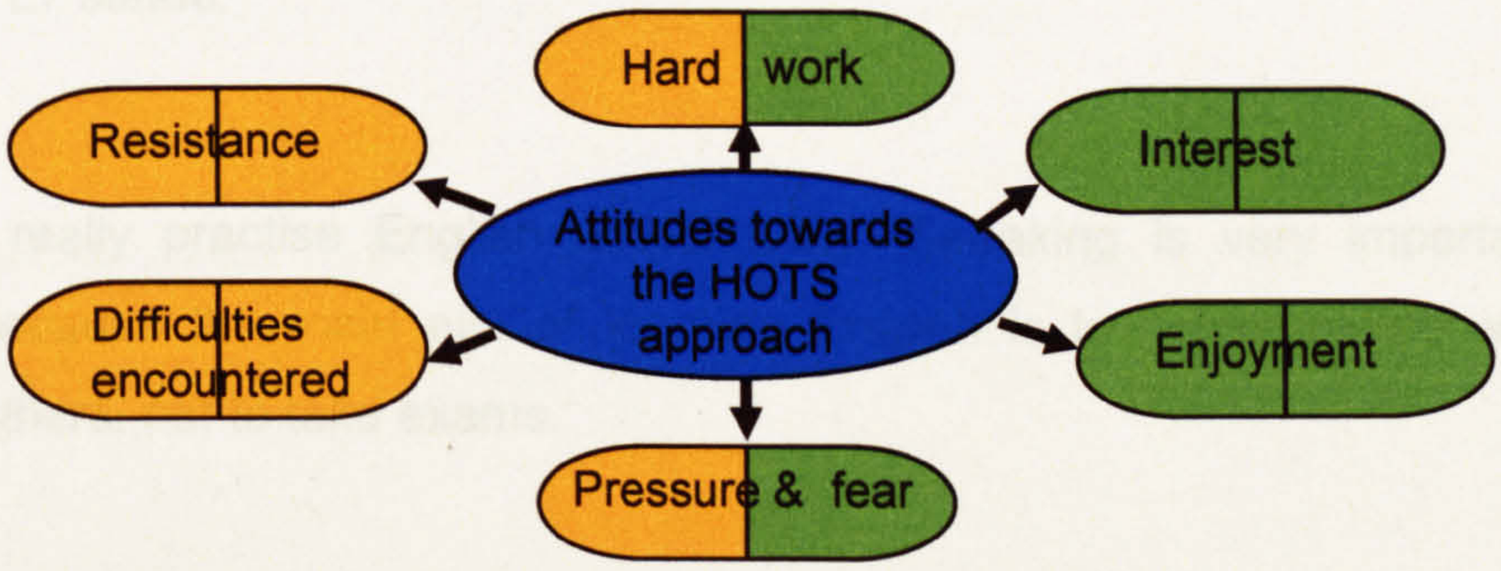


Figure 28. The attitudes towards the HOTS approach

Interest

The HOTS approach was perceived by the students to be cognitively demanding, as shown in Section 4.3.1. They considered this kind of learning process to be interesting and motivating and they participated fully, as reported by L8 and L9:

> L8: 'It is interesting to figure out the answers on my own.'
> L9: 'If thinking skills are applied to the class, the learning atmosphere will be activated.'

As the post-test questionnaire data (Section 3, No. 6) revealed, 72.5% of the students in the innovation class considered higher-order questions interesting. However, the post-test questionnaire data (Section 3, No. 7) also showed that only 57.5% of the students liked to answer these higher-order questions in English. Interestingly, this type of question required more effort and challenge in the use of the target language and as such was not universally popular.

Enjoyment

It was also found that ten of the sixteen interviewees (62%) appreciated that this innovation provided a basis for using the language to communicate. It was more practical, as L7 stated:

> L7: 'I really practise English authentically. Speaking is very important because the main aim of learning English is to communicate with others, not to take exams.'

Interviewees claimed that the discussion prompted by the tasks was like chatting in their daily lives to express ideas. They felt free to express their opinions without the pressure of being right or wrong, which enabled them to enjoy the talk more, as L9 claimed:

L9: 'I can express any opinion I want. I feel good.'

It appears that students had positive attitudes while using the target language to communicate. In addition, authentic use of language could possibly equip learners with more ability to apply the language learned outside the classroom, such as to communicate with foreigners.

Pressure and fear

Comments from eight of the sixteen interviewees related to pressure and fear, which was due to three main reasons reported by interviewees. First, students were not used to thinking and speaking in English, as stated below by L21:

L21: 'I felt the conflict towards the teaching method [the HOTS approach] used. I was worried whether I could talk in English.'

Second, the topic was sometimes unfamiliar. Interviewees claimed that when they were not familiar with the topic, it was difficult for them to think. Thus, anxiety increased, as claimed by L10:

L10: 'Now the level of pressure depends on the topic.'

The third reason relates to lower speaking proficiency. Students with lower speaking proficiency had greater fear and lower confidence. L16 stated:

L16: 'If I need to talk to them with my broken English, I would feel really painful.'

To complicate the findings, four out of eight interviewees held a positive attitude towards this pressure. Students appreciated it because they felt it supported their learning, as L10 reported:

> L10: 'I improve with that pressure.'

Most importantly, they felt they were able to overcome the pressure through the innovation as they gained in confidence and courage to talk. L12 claimed:

> L12: 'I felt very afraid in the beginning. It seemed there was always pressure in class...I did not pay any attention to the pressure. I just kept discussing and discussing. Later on I had the courage to talk, so the pressure was relieved.'

It appears that the pressure which occurred in this innovation either benefited students in their learning or put them off. In the findings from the video data, students' overall speaking proficiency was enhanced. Thus, it could be inferred that the majority of the students benefited from the pressure, rather than being put off even if they did not perceive it. .

Hard work

Five of the sixteen interviewees valued this innovation as hard work, as it was cognitively demanding for them to translate Chinese answers into English, as reported below:

> L17: 'It is cognitively demanding when translating Chinese answers into English.
> L16: It is very tiring.
> L17: Yes, then you feel down.'

In contrast, for those like L22 below, who gained a sense of achievement, hard work was interpreted as a positive comment: it benefited the learning process.

L22: 'Now I try, no matter how hard it is. I just say it out...I feel I have improved.'

The findings relating to hard work indicate that students may have been pushed out of their comfort zone where they found it hard to figure out the answers and the language needed.

Resistance

Four out of sixteen interviewees (1/4) claimed that they preferred the previous teaching approach similar to grammar-translation as it was easier and less demanding for them. L18 stated:

L18: 'I suffered a lot of pressure from it. I prefer the previous teaching method.'

The negative attitudes towards the HOTS approach reveals that it was not universally popular. It implies that either a longer time is required for students to adopt to such higher-order thinking in L2 classrooms or a mixture of pedagogies and other types of thinking tasks need to be employed in order to meet students' demand.

Difficulty encountered

Being unfamiliar with the topic was considered an impediment to students' thinking and discussion. Four interviewees claimed that when encountering an unfamiliar topic they were not able to figure out the Chinese answers, never mind the English, as L12 stated below. The finding suggests that *topic familiarity* is important when conducting the HOTS approach. Unless familiar with the topic, even students who are capable of dealing with higher-order questions will not be able to activate their higher-order thinking skills.

L12: 'It is strongly relating to our background knowledge when we could not figure out the answers. The tasks all require thinking. The difficulty encountered at the moment is the unfamiliarity with the topic.'

Inability to follow was another difficulty mentioned by two interviewees. Students with lower listening comprehension proficiency described how it was a problem for them to understand higher-order questions. Without understanding the questions, it was impossible for them to think, as L14 claimed below. This indicates that there was an impact on those with low listening comprehension.

L14: 'If there are some words I do not understand, then I will not be able to comprehend the question. Sometimes I did not understand because the teacher spoke too fast'.

In addition to students' attitudes, some issues regarding to this innovation were raised. *Insufficient learning* was the first issue raised by two *interviewees.* They claimed that learning opportunities were missed due to the fact that part of the time for teaching grammar, texts, and vocabulary was replaced by speaking practice, as L8 stated below. This result suggests that students expected the teacher to teach the vocabulary in class so that they could employ it. It also implies that answering higher-order questions demands a wider range of vocabulary items.

L8: 'It is very stressful when there is insufficient vocabulary [when speaking in English]. But I did not recite more words [this semester] due to the fact that we spent more time on practising speaking.'

The results from the questionnaire (Section 4, No.6) show that 72.5% of the students agreed they would like to learn more vocabulary and grammar. Questionnaire data (Section 4, No.7) further reveal that 75% of the students preferred to have more opportunities for practising speaking. These two findings might seem rather

contradictory due to the fact that there were only two English lessons per week. However, it seems that this innovation had an impact on students' awareness of their capacity regarding word storage and grammar knowledge; thus, they felt a need to learn more vocabulary and grammar. The findings also indicate that students felt the need to practise speaking.

Interviewees had different opinions towards the preferred forum for speaking English. Some favoured talking in groups because they felt the anxiety was lower and could talk more freely, while the others felt they could express their opinions either in groups or in front of the class: they could adapt to different talk contexts. As L12 reported:

> L12: 'I don't mind talking in front of the class. I can do it either in groups or in front of the class.'

However, questionnaire data (Section 4. No. 5) showed that 80% of the students preferred to express their ideas in groups, rather than in front of the class. This finding indicates that students had higher anxiety when talking in a teacher-led setting and group discussion was better for practising speaking.

Five interviewees reported that the length of the innovation was insufficient. If the innovation had been prolonged, they would have improved more. One of the students, L9, had a speaking ability which was much lower compared with the other two group members at the beginning of the innovation. After 12 weeks of innovation, he had improved a lot. He claimed that this innovation was not long enough for him to fully develop his speaking ability. Had the innovation been lengthened, he would have reached the same speaking level as the other two group members, as he states below. This result suggests that some students with a lower speaking proficiency may require more time to fully develop their speaking ability with the HOTS approach.

L9: 'Time was too short, something [speaking ability] hasn't been developed... I have a feeling that I could reach their level if we kept practising this way.'

4.4.2 Perceptions of the effect on students' performance

This section first reports the individual study group member's perceptions of progression from a chronological perspective observed from the interview data. A clearer picture of students' learning process and improvement can be drawn.

L7: 'After the innovation, there is not much difference with my higher-order thinking. However, it is obvious that in the beginning of the innovation it took me around two or three minutes to express one idea or one sentence, and now I just need to spend two seconds to get my idea across.'

L8: 'I feel that my thinking is getting faster and the speaking fluency increases. My confidence allows me to talk more. In the beginning of the innovation, I needed to figure out the answers to the questions in Chinese and translated the answer into English, and then I could talk. Now I can say whatever I want. The speed of speaking English is just like I think in Chinese and talk in Chinese. I now can think in Chinese and speak in English simultaneously.'

L9: 'My English speaking becomes more fluent and I think faster than before. In the beginning of the innovation, I could figure out the answers in Chinese, but I often encountered difficulties when it came to the translation into English. It also took me a long time to prepare what I wanted to say. Now I do not need to spend so much time. While the other two group members were talking, I can handle it gradually. I now become very attentive to their talk. On the one hand, I am listening to their conversation carefully. On the other, I am constructing what I am going to say reflecting on their talk... When there was a certain key word I did not understand, I

felt that I could not understand completely. Just like doing a jigsaw, one piece was always missing and I could not understand what they were talking about. Now I would say 'oh oh oh'. I understand better. My listening and reading comprehension are improving...Among the group members, my level is lower than the other two. It is a pity that time is too short. Sometimes I wanted to join the conversation, but I did not figure out the sentence in time. I have a feeling that I can reach their level if the innovation lengthens.'

L10: 'After the innovation, I do not feel my fluency has improved but the amount of talk has been boosted for sure. I have more confidence and courage in giving it a try and expressing my ideas. In the beginning, I had pressure from speaking English. I would think a lot in order to construct the sentence before I talked. I would hesitate. But now I tend to feel like to talk. I do not think that much now. I just talk in English directly whenever I think of something. This innovation has also influenced my performance in art class. I notice that I elaborate more when talking about my art work and I am able to comment on others' work. I do these more often than before. The aim for me to learn English used to be passing exam. I was interested in learning English, but I would not study if there was no exam. Now I become more like an autonomous learner. I sometimes read English articles and advertisements online which I would not do before.'

L11: 'I feel that I have made a huge improvement. In the beginning, I did not dare to talk. Even though I knew the answers to the questions, I would not like to raise my hand and talk. But later on I began to raise my hand and say it... I suffered a lot of pressure, but I got used to it gradually and overcame the fear. The psychological barrier was diminished. I also did not dare to say words I did not know before, but now I am able to use substitute words to explain words which I do not know. I have much more confidence now and I also can argue with other group members...

This innovation has also influenced my English writing. I have an Indian net pal and I wrote him email. It often took me ages to construct one e-mail in English. Now I can complete one in a very short time.'

L12: 'Last semester I didn't dare to speak English. I did not want to talk because I did not have the chance. In the beginning of the innovation, I had fear and felt that there was pressure in class... The speed of my thinking was fairly slow and I needed to think for a long time. I was afraid I could not speak out the answers or could not speak well. I also worried about my grammar and using wrong words. Sometimes, I could not figure it out, then my brain became a blank. My CPU (brain) ran slowly and sometimes it froze... I did not speak a lot until once we discussed a topic about dogs which I'm interested in. That was my turning point. I suddenly talked far more than before, and my confidence increased. Now I am not afraid of making mistakes. I just talk and talk, and the speaking fluency improves. I feel that my CPU has upgraded from 286 to Pentium 4 ... I even MSN my friends in English and found myself foreign net pals in order to gain opportunities to practise English. I didn't dare to talk to foreigners before because I did not know what to say and how to say it in English and I was afraid of making mistakes. Now talking to foreigners is not a big deal for me anymore.'

These findings indicate that the students had clearly perceived their progress during the learning process and were able to elaborate about it. They knew how their thinking and speaking were changing and the effort they devoted. This triangulated with their individual IELTS scores and speaking development discussed in Section 4.2.3. It was such cognitive training to boost the learning ability.

Having examined the individual perceptions on learning performance, it is now important to provide the overall students' perceptions of the effect on their

performance according to the interview data. Figure 29 reveals a summary of the overall effects found, namely intellect and psychology, thinking, speaking, other skills and subjects and life, ranging from the immediate impact to a wider influence on students' performance. Students' perceptions on speaking and thinking performance had been discussed in Section 4.2.2, 4.3.2 and 4.3.4. Therefore, this section will look at the effects on intellect and psychology, other skills and subjects, and life.

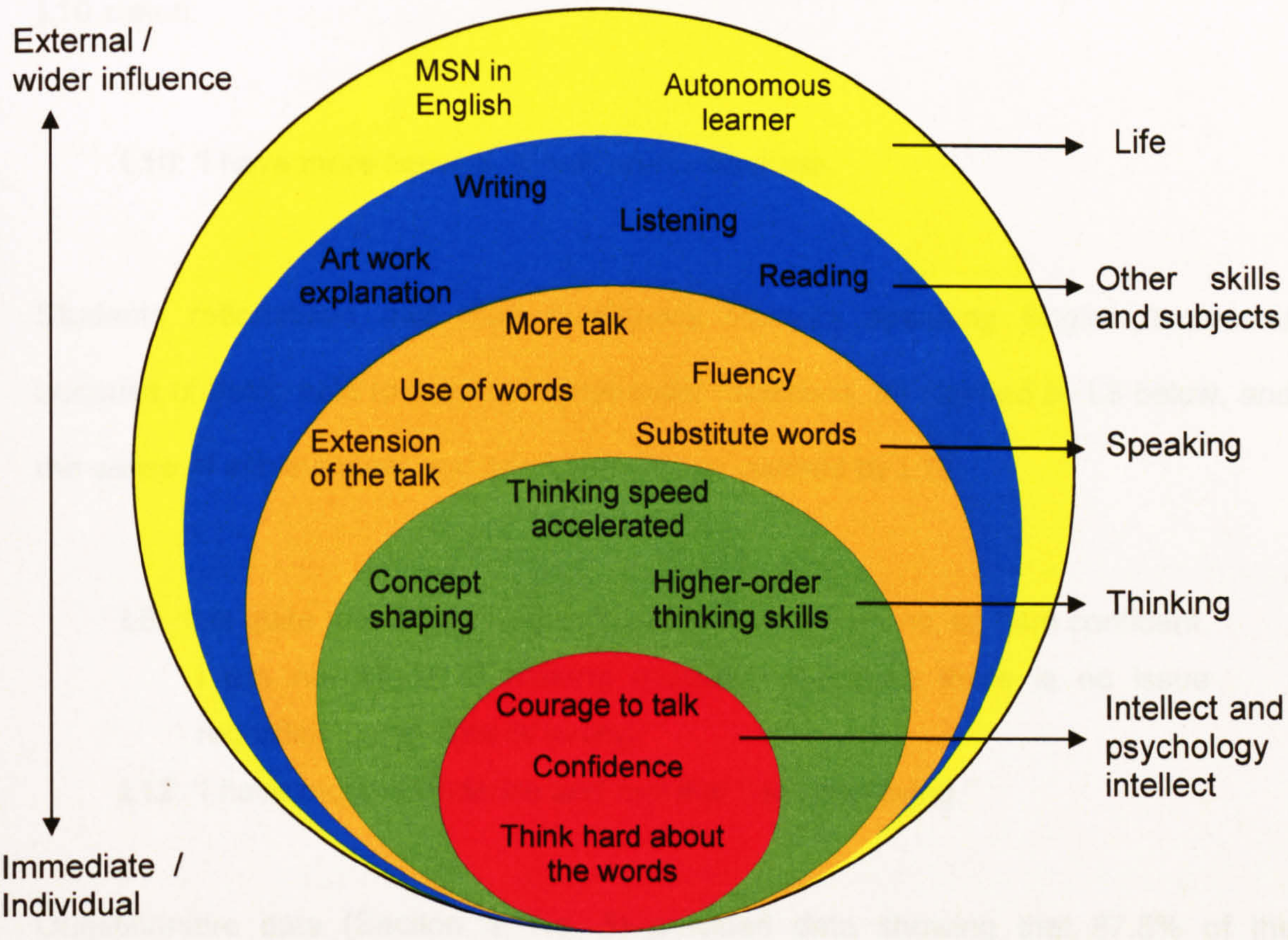


Figure 29. Perceptions of the effect of the HOTS approach on students' performance

Intellect and psychology

It is interesting to note that the HOTS approach was perceived by students to have an effect on intellect and psychology. Nine interviewees claimed they had to *think hard* in this innovation, which they rarely did before, as L17 claimed:

L17: 'My brain would go blank with Grammar-Translation Method. Now I force myself to think.'

Courage was another important improvement students noted. As Section 4.4.1 mentioned, students had great fear and anxiety towards speaking English. It was not easy to conquer such a psychological barrier. However, students argued that they gained more courage to talk and were able to take risk in terms of language use, as L10 stated:

L10: 'I have more courage to talk and to take risk.'

Students rationalised that their *confidence* towards speaking English increased because of being able to discuss higher-order questions, as claimed by L8 below, and the sense of achievement and improvement, as claimed by L12.

L8: 'It is quite relaxing to discuss higher-order questions, so I am confident. I am not afraid of making mistakes...because there is no issue regarding being right or wrong.'
L12: 'I have more confidence and feel that I am improving.'

Questionnaire data (Section 4, No. 3) provided data showing that 87.5% of the students agreed their confidence in expressing ideas in English had increased. They would give it a try, even though they did not know certain grammar or vocabulary needed. The results indicated this innovation can help to build up students' confidence in using English to communicate.

It is very positive that students considered psychological and intellectual perspectives as a part of their speaking improvement. These findings indicate that students' ability in taking risks might have enhanced.

Other skills and subjects

The effect of the HOTS approach on other subjects and skills, such as writing, listening and reading was also commented on by three interviewees. L11 reported that she noticed her English writing was much faster than before:

> L11: 'It used to take me a long time to compose one email in English. Now I can write one very quickly. I can do it more directly and do not need to think for a long time'.

L9 was aware of the improvement in his listening and reading comprehension:

> L9: "I used to feel that a little piece of information missing [when listening to others]. It was just like playing jigsaw; I sometimes missed one piece, so I did not quite understand what it meant. Now I feel 'oh, oh, oh, I see'... My listening and reading comprehension is getting better."

In addition, L10 claimed that her ability to argue and reason in other subject areas was enhanced.

> L10: 'I can talk more about art work in art class.'

These findings suggest that the use of argument-based reasoning skills and the language use may have transferred to other skills and subjects.

Life

It was also very surprising that students perceived that this 12-week innovation had influenced their daily lives. One student, L10, stated she was more aware of her autonomous learning and felt this was due to the innovation:

> L10: 'Now I become an autonomous learner... I would get on-line to read English articles and advertisements which I would not do before.'

Two students claimed that they had gained the courage to MSN in English. One of them, L12, said that he dared not use MSN in English before, especially to foreigners. He said:

> L12: 'I met a foreign net pal on line recently. I rarely talked to foreigners before. This time I did it on purpose in order to let myself have a chance to use English.'

These findings indicate that the HOTS approach provides greater opportunities to enable students to further take their learning responsibility beyond the classroom learning and apply the language learned in class to their lives.

4.4.3 Evaluation on innovation tools

The innovation tools used in this innovation were the thinking tasks, probing and wait-time. Questionnaire data were analysed to give students' evaluation on these tools.

Figure 30 illustrates the descriptive analysis of students' attitude towards the thinking tasks (post-test questionnaire, Section 4, Q8-13). More than half of the students considered '5Ws' and 'Make Up A Story' to be the two most intellectually challenging tasks which helped them most to express their ideas more logically and to develop thinking skills. '5Ws' and 'Guess What I Say' were seen as offering most opportunities for talking. 'Guess What I Say' was the students' favourite and the one they felt most confidence with. Interestingly, the findings reveal that 'Odd One Out' did not receive much attention in this innovation, although it is described as a 'fun' task in Higgins and Baumfield (2001). This task was also evaluated as offering the least speaking opportunities and students had lowest confidence with it among the four tasks. These results indicate that a successful thinking task should be cognitively challenging which can help to develop thinking, provide more speaking opportunities which is linguistically demanding yet manageable, and entertaining.

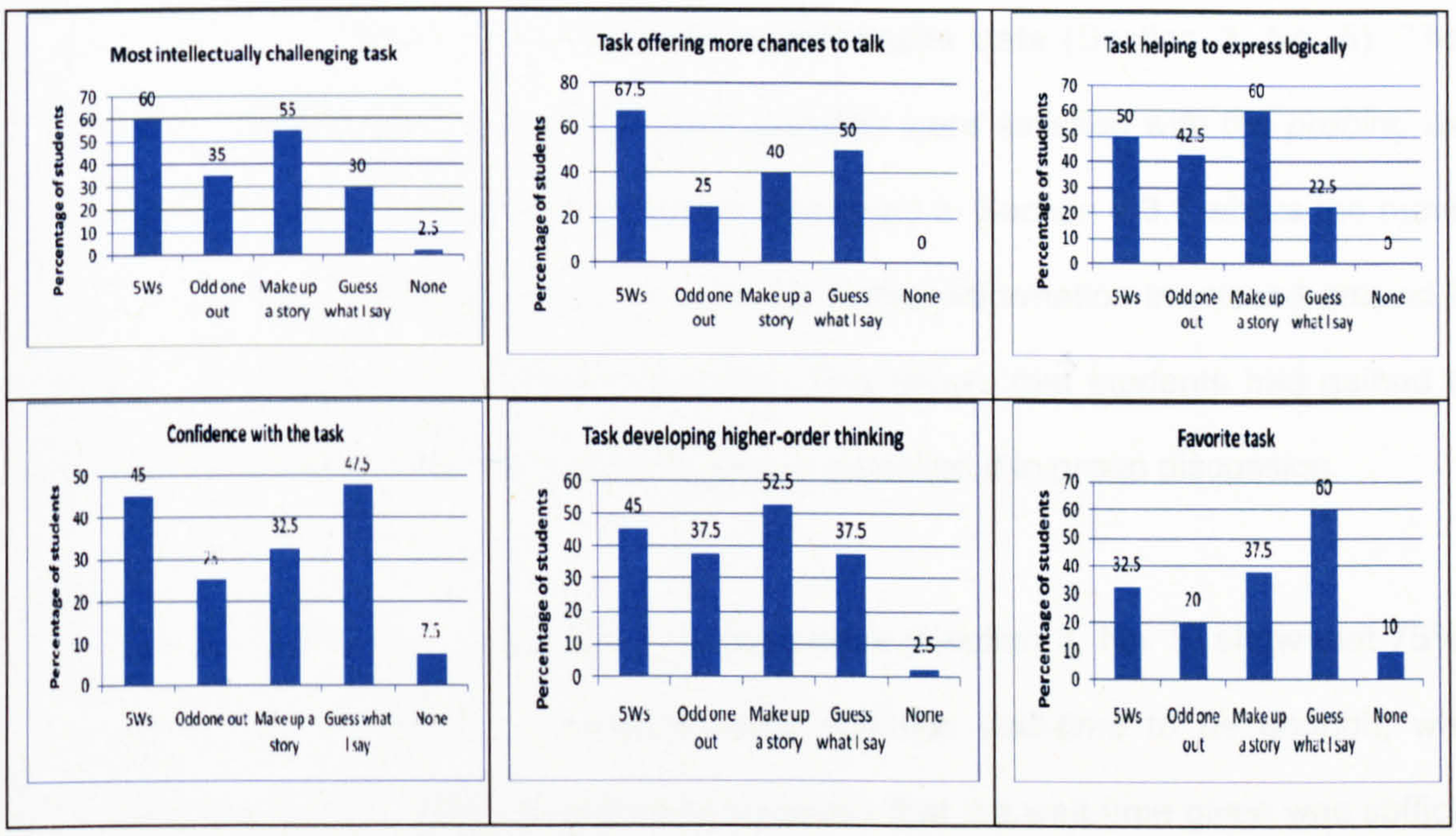


Figure 30. Descriptive analysis of attitudes towards the thinking tasks

As discussed in Section 1.5.1, different thinking tasks pose different linguistic and cognitive demands. Students evaluated the degree of linguistic and cognitive demands among the four tasks conducted, by placing them in the quadrants of linguistic and cognitive demand, as shown in Figure 31.

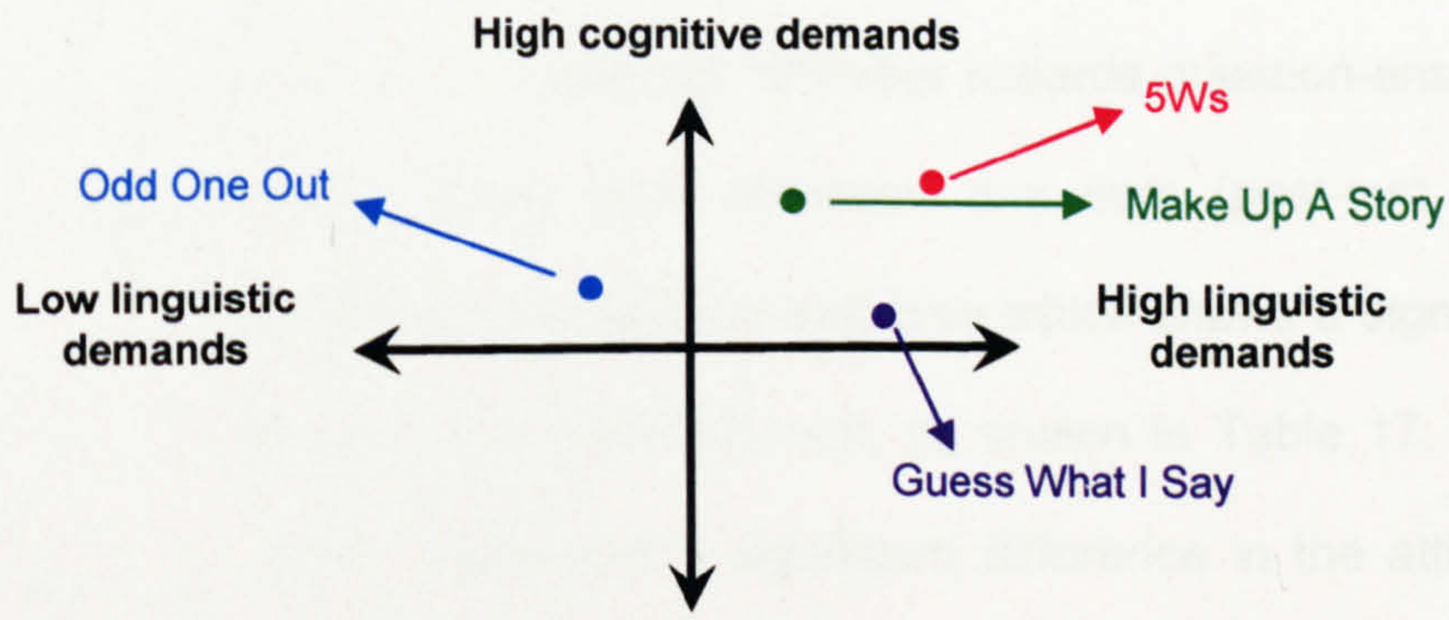


Figure 31. Quadrants of linguistic and cognitive demands of the thinking tasks

95% of the students considered *probing* to be sufficient in class, which means almost all the students agreed that the teacher asked further related questions to elicit longer

utterances, as shown in the post-test questionnaire data (Section 3, No. 5). These findings indicate that the majority of the students were satisfied with the *probing* used in this innovation. In addition, the results discussed in Section 4.3.3 where the number of higher-order questions asked to probe for further information increased around two times at both post- and delayed post-test. This shows that students had gained this skill in a teacher-led setting and were able to modelled it in group discussion.

Descriptive data from the post-test questionnaire (Section 3, No. 3) show that 75% of the students at post-test considered *post question wait-time* to be enough, which indicates that the majority of students perceived that the wait-time given was sufficient for them to figure out the answers after the teacher posed a higher-order question. In addition, post-test questionnaire data (Section 3, No.4) demonstrated that 77.5% of the students considered the *post response wait-time* sufficient, indicating that most students consider they received enough time to think about what others had said and to respond to it.

Table 16 presents the overall result of One-way ANCOVA, P= .518, which did not show a significant difference in students' attitudes towards question-answer behaviour across the two classes. There was, however, one item (post-test questionnaire, Section 3, No. 3) regarding post-question wait-time which shows a significant P value of .006 ($p < .01$) with One-way ANCOVA test, as shown in Table 17. This suggests that, after the innovation, there was a significant difference in the attitudes towards post-question wait-time given between the two classes. Table 18 presents the mean scores of post-question wait-time as 2.25 at pre-test and 2.85 at post-test in the innovation class, while for the comparison class it was 2.31 at pre-test, and 2.18 at post-test. The figures demonstrate that students in both the innovation and comparison classes considered wait-time given after the teacher posed a higher-order question was sufficient. 97% of students in the comparison class felt more satisfied

with the wait-time given. In contrast, students in the innovation class felt the time constraint during this innovation.

In addition, Table 19 shows there was a significant difference in the post-question wait-time in the innovation class with the analysis of Paired-Samples T Test (P= .009, p< .01).

Table 16. Analysis of One-way ANCOVA on the comparison of question-answer behaviour

Homogeneity of with-in regression of One-way ANCOVA

Some of Variation	SS	DF	MS	F	P
Between groups x covariant	.120	1	.120		
Error	44.261	66	.671	.179	.674

The analysis summary of One-way ANCOVA

Some of variation	SS	DF	MS	F	P
Covariant	.132	1	.132		
Between Groups	.280	1	.280	.423	**.518**
Within group	44.381	67	.662		
Total	558.704				

Table 17. Analysis of One-way ANCOVA on the comparison of post question wait-time

Homogeneity of with-in regression of One-way ANCOVA

Some of Variation	SS	DF	MS	F	P
Between groups x covariant	.417	1	.417		
Error	67.511	68	.993	.420	.519

The analysis summary of One-way ANCOVA

Some of variation	SS	DF	MS	F	P
Covariant	.047	1	.047		
Between Groups	7.836	1	7.836	7.959	**.006**
Within group	67.928	69	.984		
Total	546.000				

Table 18. Mean scores of post-question wait-time

		Pre-	Post-
Innovation	N	40	40
	Mean	**2.25**	**2.85**
	Std. Deviation	.89872	1.00128
Comparison	N	32	32
	Mean	**2.31**	**2.18**
	Std. Deviation	.89578	.96512

Table 19. Paired-Samples T Test on the perceptions of post-question wait-time in the innovation class

	Paired Differences					t	df	Sig. (2-tailed)
	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference		Mean	Std. Deviation	Std. Error Mean
	Lower	Upper	Lower	Upper	Lower	Upper	Lower	Upper
Innovation	-.60000	1.37375	.21721	-1.03935	-.16065	-2.762	39	**.009**

In brief, the descriptive results show that, although the majority of students in the innovation class were satisfied with the wait-time given, the ANCOVA test and Pair-Samples T Test both demonstrated a statistical impact on the post-question wait-time for them. This indicates that students made more effort in thinking in terms of figuring out and delivering the answers.

4.4.4 Summary

Overall, the majority of the students had a positive attitude towards the HOTS approach, while a quarter of the students saw this approach as hard work, pressure and difficulties encountered like low listening comprehension and thus, resisted it. To conduct the HOTS approach, topic familiarity is an important element in allowing students to use high cognition. However, some students perceived a transfer of higher-order thinking skills and the target language and most importantly, students were able to break through the psychological barriers in terms of speaking English and thinking.

Results from the ANCOVA tests showed a significant difference between the two classes regarding the importance of higher-order thinking skills in the L2 classroom: the comparison class rated these skills a lot more important than the innovation class. This implied that the comparison class may have experienced the advantage of thinking tasks during data collection phases. Within the innovation class itself, the results of Paired-Samples T Tests showed that students perceived a significant difference towards the importance of lower-order thinking skills: they did not consider

lower-order thinking skills as important as before. Their perception of the importance of higher-order thinking skills remained approximately the same after the innovation finished.

The majority of participants were satisfied with the probing and the thinking tasks used, except 'Odd One Out' which was considered the least interesting with the least speaking opportunities among the four tasks. This shows that the features of interest and speaking opportunity were important to students while practicing speaking. Both ANCOVA and Paired-Samples T Test show that there was a statistically significant impact on post-question wait-time in the innovation class: around one quarter of the students in the innovation class felt there were time constraint with the HOTS approach, while 97% of those in the comparison class were satisfied with the wait-time given with a teaching method similar to grammar-translation. Finally, students in the innovation class considered group discussion to be a better setting for practice speaking.

CHAPTER 5 - DISCUSSION

5.1 Introduction

The main aim of this study was to examine the practicability of the Higher-Order Thinking Skills (HOTS) approach in a Taiwanese university L2 classroom by investigating its effect on:

- students' speaking performance;
- students' thinking performance; and
- students' perceptions and attitudes.

Overall, the findings provide strong evidence to support the idea that HOTS is useful and workable: evidence collected has shown it to enhance L2 speaking proficiency and thinking performance with the majority of students having positive attitudes about its implementation.

Based on the framework of interpretation and discussion, as shown in figure 10 (Section 3.8), this chapter will bring all the findings together for a discussion. Thus, the research questions will be answered by discussing three main themes that emerged from the result: namely contexts for socio-cognitive conflict, high cognitive interactive talk and performance outcomes. Multiple sources of data are used to inform the discussion of each theme. The dynamic between these three elements is illustrated in Figure 32. Socio-cognitive conflict is created in the learning contexts which combine cognitive challenge, opportunities to practice speaking, authentic scenarios for discussion and interaction supported by the modelling conducted in a teacher-led setting. High cognitive interactive talk that resulted has a positive impact on the performance outcomes, namely L2 speaking, thinking, learning behaviours and attitudes. Both the contexts and performance outcomes can be seen to address common speaking problems observed in L2 classrooms (e.g. Yang 1992; Shou 1995;

the exploratory study) and identified by the Taiwanese Curriculum Developing Center of Higher Education (National Yunlin University of Science & technology 2007).

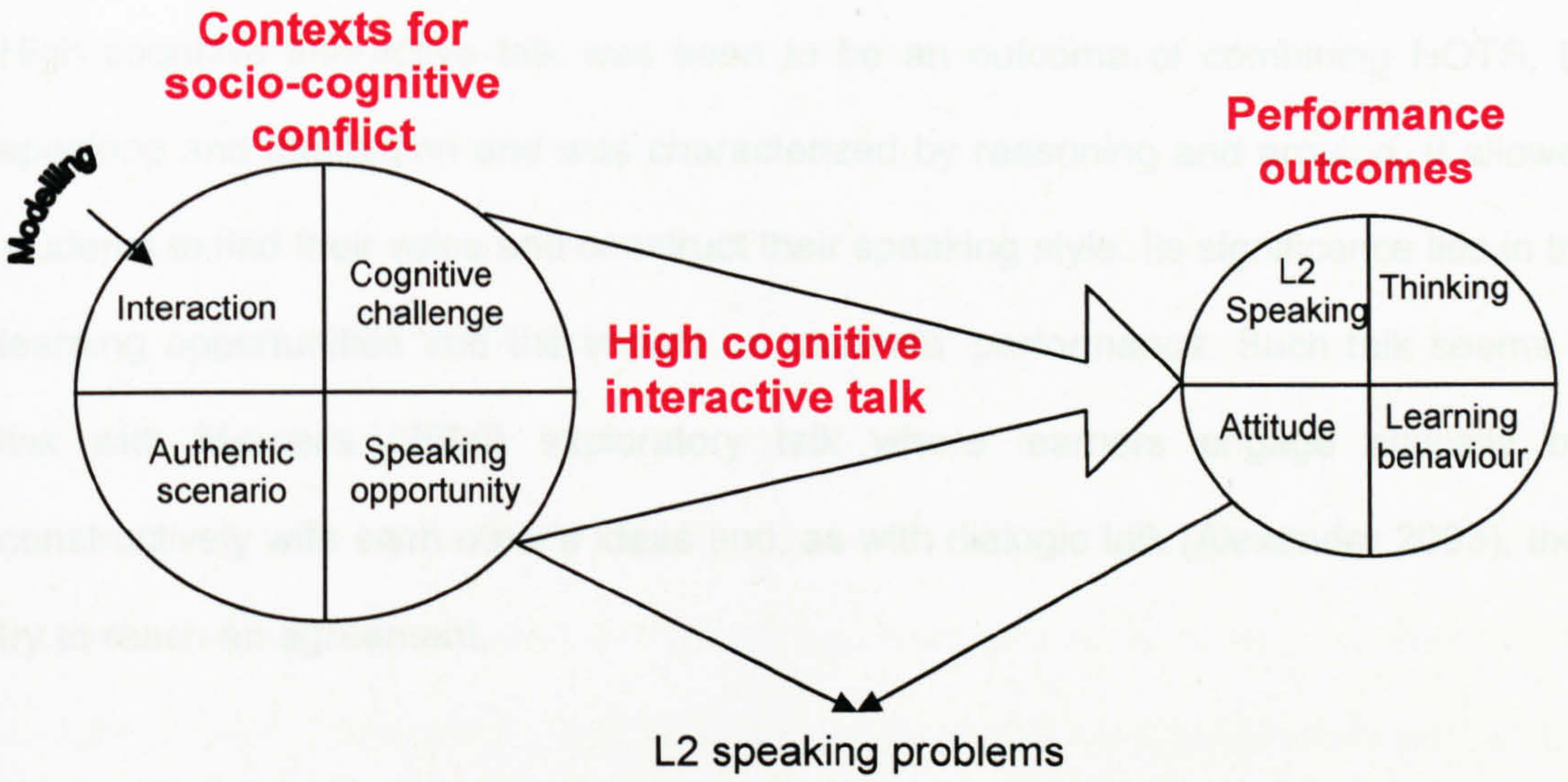


Figure 32. The overall structure for discussion

The first theme, contexts for socio-cognitive conflict, will be discussed to illustrate the types of situation created for learning. This will help a better understanding of the type of talk formed in this study and therefore the impact on students' performance. The design of the thinking tasks aimed to create contexts where students could express different opinions and generate conflict. This provided students with the opportunity to apply interaction skills to reason and argue in order to reach a consensus. Such conflict has been seen to prompt high cognitive interaction (Mugny and Doise 1978).

The speaking opportunities allowed students to communicate, control the topic and follow threads of argument pertinent to them. This *high cognitive interactive talk* can be seen to be more like real life talk characterized by unpredictability and uncertainty. It can be seen as a challenge for the students both in terms of the language and the content. The contexts created arguably tackle L2 speaking problems, for example, a lack of daily life context (Yang 1992; Shou 1995) and insufficient speaking opportunities seen in the exploratory study. This change in the type of talk used by

students was the core achievement in this study and the focus of the second theme for discussion.

High cognitive interactive talk was seen to be an outcome of combining HOTS, L2 speaking and interaction and was characterized by reasoning and arguing. It allowed students to find their voice and construct their speaking style. Its significance lies in the learning opportunities and the impact on students' performance. Such talk seems to link with Mercer's (2000) exploratory talk where learners engage critically but constructively with each other's ideas and, as with dialogic talk (Alexander 2003), they try to reach an agreement.

The extent to which high cognitive interactive talk influenced students' performance will be examined in the third theme exploring the outcomes of the innovation: impact on L2 speaking and thinking, learning behaviours and attitudes. The interaction opportunities created by the HOTS were seen to enhance oral discourse and cognition. In addition, students' disposition towards risk taking was enhanced which led to L2 improvement. Cognitive challenge required students to think proactively to solve conflicts or problems encountered during the discussion. It also had an impact on students with low listening comprehension. Surprisingly, the majority of students held positive attitudes towards the approach. Authentic scenarios, on the other hand, created greater chance for L2 transfer. The outcomes gained combine to address educational goals identified by the Taiwanese education system (National Yunlin University of Science & Technology 2007).

5.2 Contexts for socio-cognitive conflict

Thinking skills activities in association with a social constructivist approach (Watson 2000) allow the occurrence of socio-cognitive conflict (Mugny and Doise 1978; Bearison et al. 1986). This was confirmed in the current study through the use of video

and interview data and it was described by students as 'spark'. Such conflict was characterized by interaction, divergent opinions, argument and high cognitive thinking. The evidence shows the design of the thinking tasks required students to reach a consensus when conflict occurred. Importantly, when divergent ideas were expressed in the comparison class, it did not lead to an argument. The findings here are crucial as it means there was something about the HOTS approach which not only helped in creating the conflict, but also supported students in talking through issues arising and then moving towards a consensus. In addition, it is significant because these conflicts led to the occurrence of high cognitive peer interactions (Mugny and Doise 1978; Bearison et al. 1986; Mercer 2000) which has been shown to have a positive impact on L2 learning (Swain and Lapkin 2002). Without the conflicts, impact on student achievement would arguably not be as positive. The following will examine these distinctive contexts: exploring the design and implementation of the thinking tasks, how they created contexts for socio-cognitive conflict and the patterns of interaction that characterized the high cognitive interactive talk which resulted.

5.2.1 Design and implementation of thinking tasks

The design and implementation of the thinking tasks used as part of the HOTS approach affected the learning contexts created, thus the outcomes. The tasks used in the current study were arguably effective (Section 5.4). Therefore, it is important to consider what characteristics a thinking task should have: the questionnaire and interview data indicated tasks should be cognitively challenging, linguistically demanding, yet manageable and enjoyable. Students reported they needed to think about tasks, otherwise they would have nothing to say. The topics were chosen to offer the space for students to think deeply and give opinions that could be elaborated within their linguistic knowledge and schemata which fits with Coyle's (2002) idea concerning an ideal classroom task for thinking.

Without these distinct features, a task may not attract students' interest, as was indeed the case with the 'Odd One Out' task which was evaluated as the least favourite by students and providing the least speaking opportunities among the four tasks used in this study. This finding contrasts with Higgins and Baumfield's (2001) statement that 'Odd One Out' is fun in L1 classrooms. There are three possible reasons for this. Firstly, it could be the low L2 linguistic demand where students only elaborated after they figured out the answers, as reported in the questionnaire. Secondly, the technique used might not be fully appreciated by students. For example, students were not shown how to explore a topic thoroughly in modelling stage; this would need to be considered in revising the innovation. Thirdly, the topic may have been ill conceived. For example, familiarity with a topic has been shown to be an essential factor to determine a successful implementation of the tasks (Kennedy et al. 1991). It might be the L2 context and learning culture impact on the effectiveness of Odd One Out. For these students working on Odd One Out is similar to solving a mathematic problem: students would elaborate only once they figure out the answer. Future research is needed around this particular task and L2 learning.

In this study it has been shown that the design of an effective thinking task needs to build on *topic familiarity*. The interview data shows that if participants were not familiar with a topic, they were not able to think and to elaborate. This suggests that even though students have higher-order thinking ability, it may be impeded when their schemata do not provide the related information required. This finding supports previous research relating to students' performance on thinking tasks (Stapleton 2001; Kennedy et al. 1991; Carrell 1987; Franklin 1985; Glaser 1984). Therefore, in order to avoid the use of an unfamiliar topic, it is recommended that teachers support students in choosing a topic by providing a list of topics for students to choose from. By doing this high cognitive thinking skills should have a better chance of exercising students' schemata and therefore being successful.

Moving on from design, the process of implementation is also crucial in determining the success of a task. To allow high cognition to be fully executed, wait-time was seen in this study to be a prerequisite. This correlates with the work of Tobin (1987) who found wait-time to be essential in the successful implementation of higher-order questions: higher-order questions were fundamental to the design and implementation of the thinking tasks. Reports from the questionnaires showed that 75% of students in the innovation class considered post-question wait-time to be sufficient, whereas in the comparison class this was 97%. This meant the innovation class felt more time constraints. This could be because students in the innovation class were required to answer higher-order questions which resulted in them thinking harder and therefore needing longer time for this process; whereas, without receiving the HOTS approach students in the comparison class did not feel tension during any period of delay. This perceived need for greater wait-time in the innovation class could be seen to support the idea that the HOTS approach was cognitively demanding for students.

It could also be argued that the need for wait-time depends on the classroom culture and students' L2 proficiency. Honea (1982) showed that sufficient wait-time gives students more confidence in expressing their thoughts in English, but too much wait-time can contribute to student negativity. White and Lightbown (1984) suggest the ideal wait-time in L2 classrooms is five or more seconds and Godfrey (2001) discovered that foreign students required longer for answering higher-order questions than local (L1) students due to their English proficiency. The current study found that when students were accustomed to using HOTS, they could respond more quickly. Therefore, both students' L2 proficiency and the familiarity with HOTS can affect the wait-time required. Thus, there is a complex relationship which needs to be interpreted by the teachers depending on the changing needs of the students.

During the innovation, thinking tasks were conducted both in a teacher-led setting and small groups. Different settings have been shown to impact on students' anxiety (Tsui 1996). The findings obtained from the questionnaire data confirmed this showing 80% of students preferred to express their ideas in groups rather than in front of the whole class. As Tan (2007) observed, students felt nervous in a teacher-led setting and found it hard to think critically. In a small group discussion however pressure and anxiety can be alleviated, so students feel more at ease and confident to speak and conduct deeper thinking (Tsui 1996). With such a learning context, as Thornbury (1999) states, learners are likely to be more actively involved in the learning process, rather than being passive recipients. Thus, the findings in the present study suggest that students are likely to be more attentive and motivated in their learning in small group work. However, it should be recognized that class discussion is necessary for the teacher to model to students how to interact and use the language to communicate. So again there is a need for the teacher to consider the most appropriate pedagogy and balance levels of discussion appropriately.

Recommendations for instruction and material design of thinking tasks can be drawn from the above discussion. An effective thinking task in a HOTS approach should nurture learners through a cognitively challenging yet linguistically manageable and demanding equation in which students enjoy working. Topic familiarity needs to be considered while designing a task around an unfamiliar topic can inhibit performance. Wait-time is also crucial in determining success. The length of wait-time required for thinking classrooms depends on students' L2 proficiency and the familiarity with HOTS. Tasks that encourage group discussion could be a better forum for students to think critically and explore authentic talk, however teachers also need to model what this behavior looks like and so a combination of forum is needed. If a thinking task lacks any of these core ingredients, it might not work as effectively as the results

gained in the current study. How the design and implementation of thinking tasks affected the contexts created will be scrutinized in the following section.

5.2.2 Thinking tasks as contexts for socio-cognitive conflict

High cognitive thinking was observed in the video data as a result of the interaction created by the thinking tasks. This way of generating high cognition is different from that driven by higher-order questions posed by the teacher or a handout. This shows that learners can think proactively with high cognition in a social context using thinking tasks.

Thinking tasks were considered by 60% of the students (questionnaire data) as cognitively challenging. This finding is in line with Coyle's (2002) argument that the cognitive skills involved in processing L2 language input and output are complex and demanding. The demands stemmed from both the language use and interaction: listening and concentrating to group discussion, analysing information received and simultaneously forming their own views on the topic, working out grammatical rules, understanding the norms of use and appropriateness and sequencing utterances which are meaningful and can be understood.

The interaction that occurred increased the level of cognitive challenge. As argued by Vygotsky (1978, 1986) and Lantolf and Appel (1994) in a sociocultural context, higher cognitive processes can emerge as a result of the interaction. This situation was described by students in the present study as 'brainstorming' which required them to think hard when judging others' opinions, probing and extending the topic of interest so that they could generate something to talk about and to interact with others. During the interaction, thinking speed could be seen as another challenge. As L9 said he could not processed the information simultaneously in the beginning so that he was not able to join the conversation. It was when he got used to such cognitive challenges and

was able to accelerate his thinking, he could then enter the discussion. Thinking tasks by their nature require students to devote more effort to thinking (Higgins and Baumfield 2001). Yet, it would be interesting to learn whether cognitive challenge remained as the innovation went on (Section 5.4.1).

One interesting phenomenon worthy of discussion here is that students claimed in the interview that they had high cognitive skills and could use them freely in Chinese to figure out answers to the higher-order question listed on the handout. However, when students failed to translate the answer into L2 and were not able to elaborate due to the deficiency in their English proficiency, it does not necessarily mean that students do not have high cognitive thinking. It is a matter of how to exercise these cognitive skills through the target language; without practicing in English it was much more difficult for them to operate these skills in L2. As Godfrey (2001) observed any language barrier can be an obstacle to employing HOTS successfully. It is a potential limitation of this study that students' cognitive levels were not tested in their L1 before and after the innovation, thus it is not certain the extent to which the HOTS approach had an impact on their cognition when they were using L1.

The thinking tasks offered chances for students to elaborate on their ideas, yet it was high cognitive peer interaction which built up more opportunities for speaking. The interaction created plentiful space for expanding one's own thinking and commenting on others' thoughts. As a result, further topics for discussion and greater space to talk about it were generated. Thus, it is not surprising to find in the questionnaire data that 80% of students considered greater speaking opportunities were provided with HOTS when they could extend the topic of interest. This fits with Ellis's (1990) suggestion that optimal conditions for L2 learning are characterised by the need to use the target language and opportunities to initiate and control topics. The thinking tasks used in the current study aimed to create opportunities for talk, the need to use the target

language authentically and the chance to initiate and control the topic. The finding implied the HOTS approach provides optimal learning contexts for L2 speaking.

The nature of the interaction and the ability to initiate and control topics implies real language use (Coyle 2002) characterized by unpredictability and uncertainty. This was confirmed in the video and interview data where authentic language circulates when L2 was used to express thoughts and communicate (Talebinezahd 2003). The finding implies the HOTS approach provides greater opportunities for authentic language use. The talk, unaided by the teacher, was unscripted and natural. This has been shown to be a decisive factor in enhancing L2 speaking. As Richards and Rodgers (2001) and Williams (1998) claimed, language used for authentic communication is the most effective way to learn, especially if it contains value for the students, for example, when used for expressing beliefs and ideas (Williams 1991). In the current study, some tasks were value-laden, e.g. criteria for choosing an ideal mate. Since there were fewer constraints on the outcome of the tasks, it provided an open ended context for learning which enabled students to talk more authentically.

The enhancement of L2 speaking proficiency (Section 5.4.2) supports the idea that an authentic scenario is crucial for L2 learning. According to Leith (2000) and Smith and Paterson (1998) talking for real requires learners to think deeply about how language fits together. Students in the present study become more aware of sophisticated language use by gaining insights into how language can be substituted and used for reasoning and questioning. The findings suggest that an authentic scenario facilitated L2 speaking enhancement.

Using authentic scenarios can be seen to additionally increase students' learning motivation. Students in the current study never knew what they would be asked to talk about next and therefore tackled subjects they would not have talked about otherwise.

The conversations produced were particular to the context and these interlocutors. This produced increased levels of challenge. Such unpredictability and uncertainty have been shown to increase motivation in the language learning context as observed by Rodgers (cited Thomas 2000). In addition, such authentic scenarios may promote L2 transfer. This issue will be further discussed in Section 5.4.3.

The thinking tasks encouraged peer interaction and this can be linked to theory of collaborative learning where students work together towards a common goal (Gokhale 1995). The sense of achieving a consensus facilitated the occurrence of peer interaction. Students in this study reasoned for their own value and tried to persuade others in order to come up with an answer agreed by everyone. Thus, the HOTS approach facilitated interaction that kept going until an agreement was reached.

Collaborative learning has also been shown to help students learn from each other's knowledge, skills and experiences (Johnson and Johnson 1986; Gokhale 1995). Contributing to the joint meaning-making with and for others extends one's own understanding (Wells 1999). Thus, it is possible for learners to internalize both external knowledge and critical thinking skills, and to convert them into tools for intellectual functioning. This explains results found in this study where the relationship between high cognitive peer interaction and cognitive and speaking performance was positive and the knowledge was co-constructed.

As Gokhale (1995) observed, collaborative learning fosters the development of higher-order thinking through discussion, reasoning and evaluating others' ideas. In the present study, students needed to go beyond mere statements of opinions by giving reasons for their evaluation and reflecting upon others' comments. Such active exchange of ideas within small groups promotes higher intellectual thinking and therefore corroborates the finding of Gokhale (1995) and Johnson and Johnson

(1986). This shared learning gives students opportunities to engage in high cognitive thinking and interaction (Totten et al. 1991).

To sum up, the contexts created by the thinking tasks provided an optimal L2 learning condition. It meant students had opportunities to think critically and explore authentic situations through dialogue. They could build up their confidence and participate in peer interaction. These are all critical aspects of a quality learning context for L2 speaking. The interaction affected the contexts created, like increasing the cognitive challenge level and speaking opportunities. So it is now important to explore what this interaction pattern looks like.

5.2.3 Exploring patterns of interaction

The thinking tasks were designed to provide participants with opportunities to articulate various opinions and exchange their views. This did not mean peer interaction, that is conversation among peers, would occur or increase. Video evidence showed different interaction patterns occurred in the innovation and comparison classes, as demonstrated in Figure 33. It is possible to argue that the patterns that occurred when the HOTS approach was used, in the innovation class, produced high cognitive interactive talk (Figure 34, discussed in full in the next section), whereas patterns observed in the comparison class were more in line with high cognitive monologue.

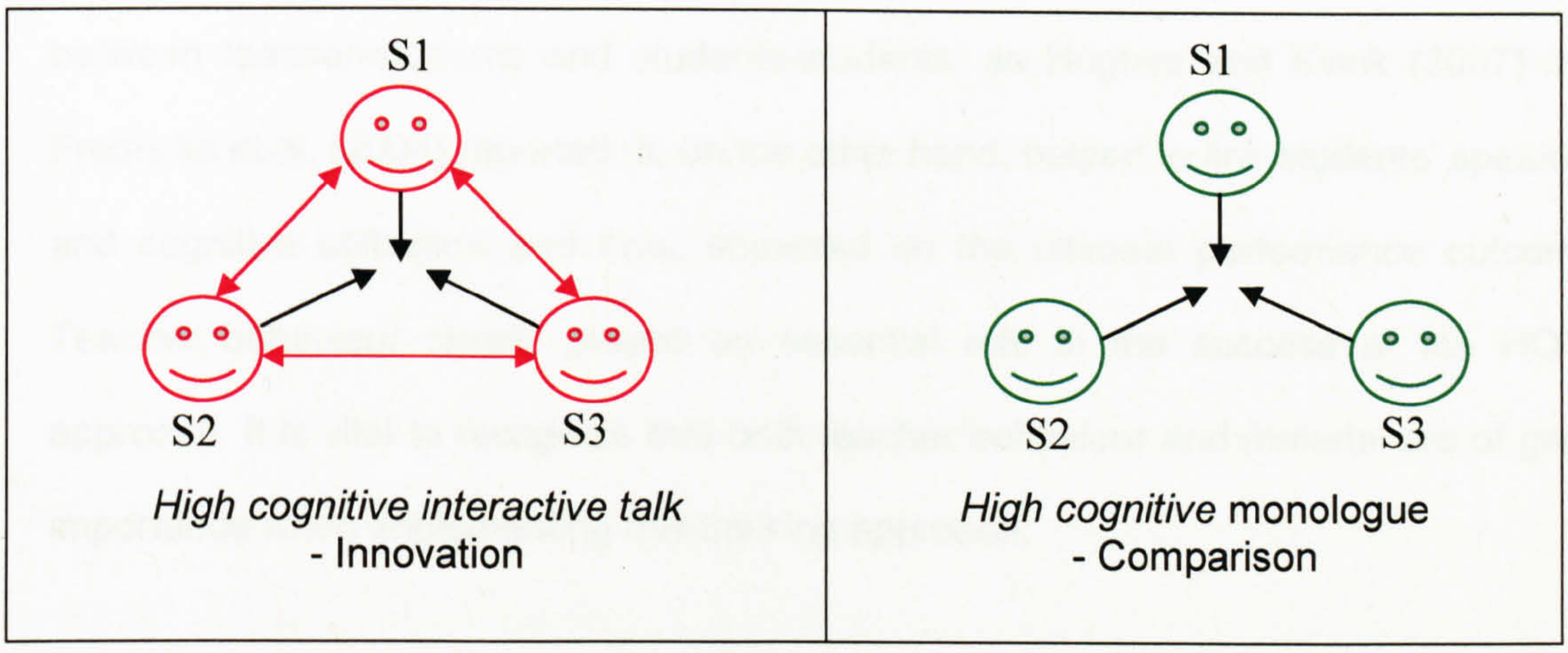


Figure 33. Patterns of peer interaction

The interaction pattern in the comparison class can be described as *high cognitive monologue* (Figure 33). The characteristics of this interaction combine a predominance of higher-order questions and students concentrating on articulating their individual thoughts in reaction to those questions. Rather than responding to each other, lack of peer interaction in the comparison class could be explained by the fact that students did not receive the support of language and techniques to interact with others. As Johnson et al. (1993) argue, just putting students in groups and asking them to work together may be insufficient to generate the kind of language and learning desired. In other contexts, to encourage students to learn together effectively, a wide range of techniques have been developed, such as providing appropriate language to use (Kagan 1994).

Modelling, using higher-order and probing follow-up questions and demonstrating the interaction skills and language use for communication, was an important factor to trigger the occurrence of interaction in this study. Such teacher behaviour mediated between students' schemata and classroom engagement: students actively used their background knowledge to participate collaboratively, self-directed the talk, persisted, paid attention to others' opinions and devoted greater amount of effort in terms of cognition and speaking. This engagement, on the one hand, mediates the relatedness

between teacher-students and students-students, as Hughes and Kwok (2007) and Fredricks et al. (2004) reported. It, on the other hand, helped to link students' speaking and cognitive utilization and thus, impacted on the ultimate performance outcome. Teacher behaviour clearly played an essential role in the success of the HOTS approach. It is vital to recognize that both teacher behaviour and material are of great importance when implementing this thinking approach.

Yang et al's (2005) study pointed out the importance of modelling to allow the interaction with high cognition to occur. The teacher models to the students, so students can imitate accordingly: to justify each others' opinion, agree or disagree with, probing and expanding the topic while answering high cognitive questions. In the present study, language use and communicative skills to probe, comment and build on others' ideas were modelled to students. This modelling arguably built up students' communicative competence (Saville-Troike 1996). As argued by MacIntyre et al. (1998), communicative competence, self-confidence and a desire to converse which this study confirmed (Section 4.4.1 and 4.4.2) are essential to motivate students' willingness to communicate, using a L2. Such modelling prepared students to be psychologically ready with language use and interactional skills for the interaction in group discussion. Thus, the modelling impacted on the interaction patterns of the innovation class.

The interaction patterns observed in the innovation class as a product of the socio-cognitive conflict had characteristics that included a predominance of higher-order questions, students concentrating on elaborating their views and commenting on others' opinions. This is indicated in Figure 33 using red arrows and fits with understandings of high cognitive peer interaction. Such interaction greatly influenced the improvement achieved. It is interesting to note that there were occasions when high cognitive interactive talk patterns stopped. These incidents were seen, for

example, when students had different interpretations of a certain word. At this point, low cognitive thinking, e.g. knowledge and comprehension, was used for communication. High cognitive interaction would only be operated once they reached the mutual understanding of the interpretation of the word.

Without the modelling, peer interaction might not occur and the dividend yielded by the HOTS approach would not be as significant as the outcomes gained in the present study. This modelling can be seen as a precondition to allow this approach to be effective. Reasons for why peer interaction is superior to non-peer interaction will be discussed through exploring high cognitive interactive talk observed in the innovation class. In the following section, it will examine the second theme of the discussion: the characteristics of high cognitive interactive talk, the learning opportunities that were created as a result and the benefit seen for students.

5.3 High cognitive interactive talk

High cognitive interactive talk was achieved in the innovation class as a result of socio-cognitive conflict when learners were engaged in the practice of HOTS, knowledge building and L2 use in a social context. High cognitive oriented dialogue was observed among students in the present study when they helped each other to scaffold talk through reasoning and arguing, and this is where thinking and speaking occurred co-constructively (Mercer 2000; Van Boxtel et al. 2000), as shown in Figure 34. The co-construction of thinking and speaking was reflected by the appearance of talking about various ideas, responding to questions, relating commenting or building on each other's opinions, elaboration of conflict and construction of reasoning and argument. That is to say, the talk and the use of high cognition rose together in an iterative process of mutual support such that more talk led to more use of high cognition which in turn led to more talk, in a virtuous circle of development. Such talk, as Coyle (2002) stresses, is spontaneous, learner-driven and focused.

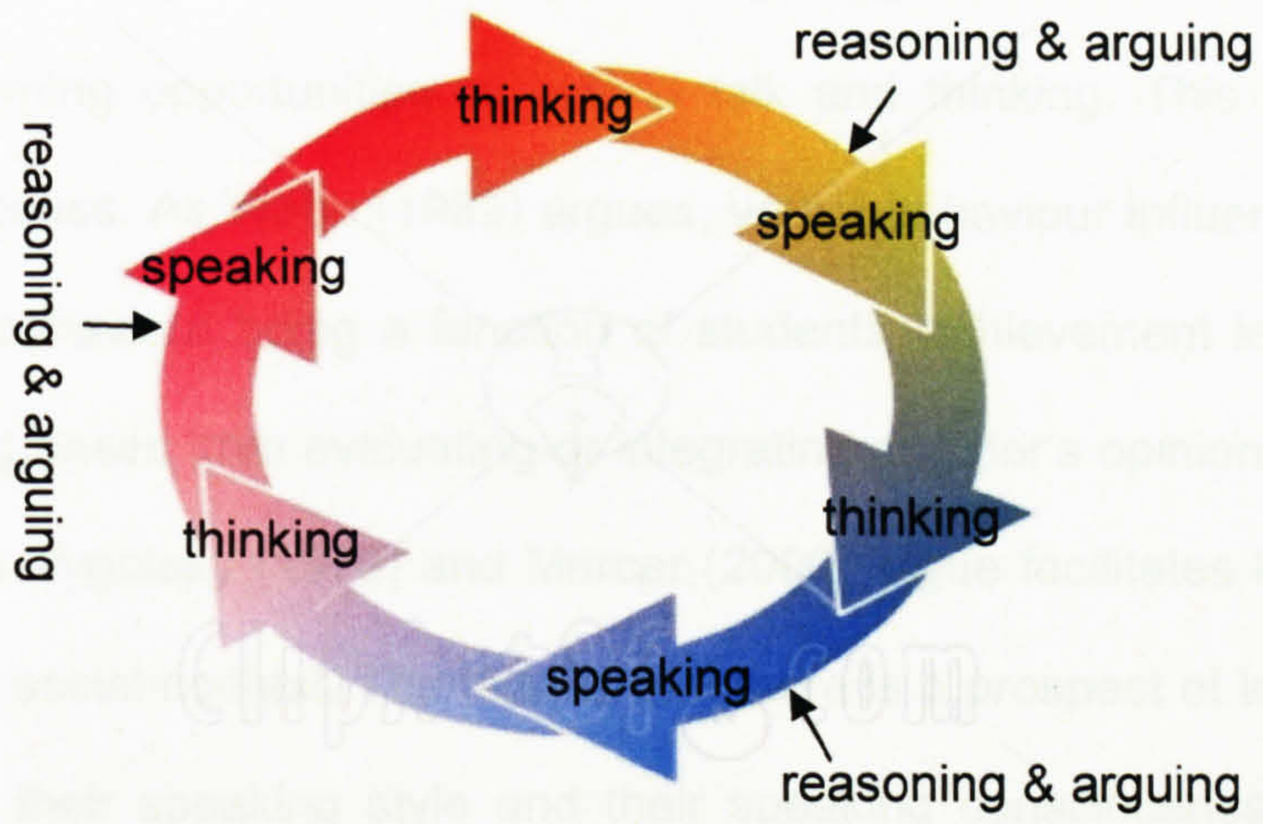


Figure 34. High cognitive interactive talk wheel

This type of talk is similar to Swain's (2000) and Swain and Lapkin's (2002) collaborative dialogue in which speakers are engaged in solving linguistic problems and building knowledge about language. Collaborative dialogue, however, focuses on the L2 linguistic domain, while high cognitive interactive talk, as conceived in this study includes both content as well as linguistic development. High cognitive interactive talk is also similar to Mercer's (2000) exploratory talk where learners engage critically but constructively with each other's ideas to reach a consensus and solve a problem. It also draws on common ideas to Alexander's (2003) dialogic talk which is collective, reciprocal, supportive, cumulative, purposeful and productive where expressing opinions, reasoning, questioning and arguing progressively build into coherent and expanding chains of thinking and speaking. However, both exploratory and dialogic talk have been observed in L1 classrooms, the talk in the present study, as with the work of Emerson and Holquist (1981), therefore needed to be conducive for both language learning and cognitive development in L2 classroom. Giving high cognitive elaborations has been found to be positively related to the explainer's achievement (Webb 1989). The language used arguably was remarkable from a linguistic perspective which contained an increasingly complex syntax (Godfrey 2001).

There are three significant advantages of *high cognitive interactive talk*. One is that it creates learning opportunities to extend talk and thinking. This is influential in the learning process. As Webb (1989) argues, verbal behaviour influences learning rather than such behaviour being a function of students' achievement level. The second is that thinking arises from evaluating or integrating another's opinion against one's own, and this, as Vygotsky (1978) and Mercer (2000) argue facilitates independent mental activity in a social context. The third is that there is a prospect of learners finding their own voice, their speaking style and their speaking consciousness which are socio-culturally acceptable (Coyle 2002). Ramanathan and Kaplan (1996b) contend that presenting a strong voice is a Western notion which is not necessarily relevant in other cultures. However, in line with Stapleton's (2001) work, this study shows Taiwanese students have individualized voices through high cognitive training. As shown in the video and interview data, students in the innovation class had personal opinions and could defend their own position. This type of talk trained students to think proactively and they constructed their own speaking style including how to reason, probe the topic, extend the talk, use substation to express ideas, etc.

In contrast to *high cognitive interactive talk*, the video data showed that the type of talk that occurred in the comparison class was *high cognitive monologue.* Students focused only on expressing their own ideas using HOTS without interacting with others in a group discussion forum. As collaborative learning rarely occurred, their talk centred on questions listed on the handout to provide reasons. The talk was therefore monologue rather than conversation. This type of talk was not dissimilar to answering the teacher's question in a teacher-led setting which could, as Dillon (1978) and Yamada (1913) have argued, make learners passive and not facilitate independent thinking.

Overall, *high cognitive interactive talk* is about voicing, thinking and communicating. It indicates that the HOTS approach not only encourages peer interaction, but also creates numerous learning opportunities, provides chances for independent thinking in a social context and the development of a speaking style. Such talk greatly influences the operation of HOTS and speaking in the learning process. The following sections will discuss the extent to which *high cognitive interactive talk* affected thinking and speaking, the third theme of the discussion, the performance outcomes.

5.4 Performance outcomes

The characteristics of the thinking tasks, as stated by O'Donnell and Dansereau (1992), influenced the type and the amount of interaction, consequently affecting the outcomes with regards to student performance. In the present study these learning contexts produced socio-cognitive conflict where high cognitive interactive talk was formed. Within these contexts high cognitive thinking skills and speaking built on each other and connected like a chain where maximum opportunities for students to communicate were created. Students' thinking and speaking were strongly affected. Such learning contexts also directly influenced students' learning behaviours, disposition towards risk-taking and awareness of the learning process. The design and implementation of the thinking tasks, the contexts created and the outcomes gained influenced students' attitudes towards and perceptions of the HOTS approach. It is now necessary to examine the impact such optimal L2 learning contexts had on students' performance: cognitive cultivation, L2 speaking development, learning behaviours and attitudes.

5.4.1 Cognitive cultivation

Thinking tasks, by their nature, provided cognitive challenge. On top of that, high cognitive peer interaction multiplied this challenge by requiring students to use high cognition proactively in conversation. This appeared to enhance students' thinking.

The results obtained from the video data showed students' cognitive performance in the innovation class was much improved when compared to the comparison class. Cognitive performance was determined by the difficulty level of HOTS in this study: the frequency of HOTS and the ways high cognition was used in a social context, for example merely for expressing opinions or interaction. *Difficulty* (Section 2.4.3) refers to the amount of effort that the learner must expend within a level of complexity to accomplish a learning objective. The analysis tool used to examine how high cognition was used in the social context was developed by the researcher; it has not been validated in the literature; thus, the validity of this finding needs to be considered. However, some triangulation of the data is gained from the questionnaire which demonstrated that 72.5% of students considered their HOTS developed.

The current study showed an improvement in the frequency of higher-order thinking from students using the HOTS approach when compared to the comparison class. This finding is consistent with those of Willson (1973), Cole and Williams (1973), Arnold et al. (1974), King (1990), and Miri et al. (2007) in L1 classrooms, as well as the research of Ayaduray and Jacobs (1997) and Godfrey (2001) in L2 classrooms. In addition, the results of the present study concur with the findings of Eisenman and Payne (1997), the Thinking Skills Review Group (2004), Higgins et al. (2005) and Burke and Williams (2008), who all conducted thinking skills programmes in L1 classrooms, concluding that thinking skills activities prove to be effective in improving learners' cognitive performance. The result indicates that the HOTS approach encourages the frequent use of high cognition.

Regarding how HOTS were used in a social context, the findings show that the innovation class frequently utilized higher cognition for communication, while the comparison class merely used it for expressing one's own thoughts rather than for peer interaction. No empirical research has been found examining cognitive

improvement by looking at both aspects. This can be argued as a more thorough inspection than showing just the frequent use of high cognition, as occurred in other research (e.g. King 1990; Godfrey 2001). The result implies that peer interaction in a social context as part of the HOTS approach can enable students to increase the difficulty levels of higher order thinking.

The difficulty of HOTS can be shown by looking at the frequency of high cognition used by students, yet crucially genuine ability lies in the context where students *use* this high cognition. Arguably the more higher-order questions are posed or the more students elaborate their answers using high cognition, the higher the frequency of HOTS. It is through the exploration of how HOTS was used in the context that this aspect can be explored. The design of the tasks allowed socio-cognitive conflicts to emerge where the conversation was boosted: when students exchanged ideas, reasoned and argued. Such rapid conversing and processing of information indicated high cognition was much more actively operated than merely when just elaborating opinions as in the patterns of interaction observed in the comparison group. This can be seen to represent a higher difficulty level. Such in-depth understanding towards cognitive development in the field of HOTS is arguably very distinctive and needs to be highlighted.

In addition, the cognitive improvement was sustainable. The findings from the delayed post-hoc video data showed students were still using HOTS for interaction after the innovation finished. This has not been found previously explored in L2 classrooms, although there are emerging L1 findings (for example, Yang et al. 2005; Zohar's 1996). The sustainability found in the current study indicates the HOTS approach is effective in developing high cognition in the L2 classroom as well as L1 classrooms and is sustainable. It is worth further exploring whether the positive impact of the HOTS

approach on thinking occurred in a social context has any influence on individual cognitive performance.

The present study confirms the significant effect of collaborative learning theory on cognitive development (Section 5.2.2). As discussed in 5.3, high cognitive interactive talk created numerous opportunities for thinking. It is now important to examine whether such interaction had a relationship with cognitive performance. Video data revealed a positive association: when high cognitive peer interaction increased, the frequency of HOTS was also enhanced. This result is in line with findings from Johnson and Johnson (1986), Gokhale (1995) and Sutherland (2006) who found students working together in a group achieved higher levels of thought and performed significantly better in critical thinking than students who worked individually. Consistent with Bloom's (1956) beliefs, many educators also claim that active mental participation encourages thinking (Long and Porter 1985; Cam 1995; Mercer 2000; McGregor 2007). Such interaction and shared learning gave students opportunities to engage in discussion, take responsibility for their own learning, and thus become critical thinkers (Totten et al. 1991). The positive relationship found implies that collaborative learning emerging in the HOTS approach facilitates cognitive development.

The results of cognitive enhancement suggest that the HOTS approach is effective in fostering thinking in the L2 setting. This is in contrast to the argument made by Atkinson (1997), Ramanathan and Kaplan (1996a, 1996b) and Fox (1994), that non-native speakers of English, in particular Asian learners are deficient in critical thinking abilities and it is difficult or impossible to teach critical thinking in L2 classrooms. The result of the present study supports Day's (2002) finding that Asian students can be trained to think critically. As Fisher (1998) claims, thinking skills can be developed through training, education and experience. In fact, the main issue is whether teachers have sought the ways to impose such high cognitive thinking on students; it is not a

matter of students' initial deficiency of critical thinking (Benesch 1999). The present study confirmed the appropriateness of using high cognitive training to cultivate students' thinking in L2.

On the other hand, the positive effect on thinking implies that students who favour the HOTS approach may have a better chance to be converted from passive, rote learners to active learners. It could be argued that students accustomed to rote learning are not necessarily unable to think proactively or learn actively; what they need is the opportunity to learn to think. This finding corroborates Littlewood (2000) and Stapleton (2002) where Asian students do not wish to be spoon-fed with facts but would like to be active and independent in learning, although their learning style is passive in class due to the educational contexts. The Taiwanese educational goal to reduce the use of rote learning and equip students with thinking skills (Ministry of Education of Taiwan 2001; National Yunlin University of Science & technology 2007) could be met somewhat by this approach.

The cognitive improvement found suggests the application of the HOTS approach might also be effective in L1 classrooms in Taiwan. As Chou (2006) found that Taiwanese university L1 learners can be trained and equipped with critical thinking skills. The literature suggested that thinking skills programs work well in western L1 classrooms (e.g. Higgins et al. 2005; Thinking Skills Review Group 2004) and have been well documented in improving oracy, cognitive performance and student achievement. It could therefore be argued that without the language barriers that exist in L2 classrooms (Godfrey 2001), the HOTS approach may be more effective in L1 classrooms. This can be an important and valuable research area in Taiwan, since cultivating students' thinking skills is an important issue in HE (National Yunlin University of Science & technology 2007).

The results further suggest that embedding thinking skills in regular language instruction enables effective cognitive development. As argued by Wenden (1997) and McGuiness (1999), integrating thinking skills in subject content may be the most effective approach for the learning of thinking. In contrast, if they are to be taught in isolation, as pointed out by Jones et al. (1987), they might not transfer across the curriculum, especially with less proficient learners.

In addition, skills related to thinking like problem solving, accelerated thinking and argument-based reasoning were seen developed in this research. The video and interview data evidence students learning to solve problems by negotiating when the problem of *word interpretation* occurred. Instead of avoiding it, students learned to solve impediments to their discussion, thus providing opportunities to revisit definitions and restructure their knowledge (Mugny and Doise 1978, 1979; Bruner 1986). As Bruner (1985) contends, when students are confronted with different interpretations in a social context, their problem solving strategies can be improved. This finding implies that the HOTS approach provides better opportunities for developing problem solving skills. This ability, as Bentley (2000) stresses, helps students to think flexibly and develop the competencies needed to face any challenges in everyday life.

The data shows students' thinking was accelerated. This is in agreement with Sousa (2001) who argues that learners can manage a higher difficulty level of HOTS and as a result retrieve information faster; in other words the learners can process the information with high cognition, retrieve their schemata and language storage and respond instantly. This finding could be considered less reliable because the tool used to assess the thinking speed, looking at the *difficulty level* where the amount of information processed and the time consumed, has not been validated in the literature. However, the researcher would argue that in the context studied, the means used is reasonable and logical to examine the change of thinking speed. Fast thinking is an

important skill in conversing in a social context, as L9's case discussed in Section 5.2.2. It is also essential in daily life as the present world is overwhelmed with the information. This finding indicates that the HOTS approach may enable students to think more efficiently when processing the information.

The accelerated thinking indicates a potential reason why cognitive challenge, one characteristic of thinking tasks, declined as the innovation went on. When students became accustomed to higher-order thinking skills and their high cognitive performance was enhanced, the tasks appeared to no longer be as challenging as before. In addition to cognitive enhancement, increased L2 speaking ability also contributed to this decreasing challenge. As students' speaking proficiency strengthened, the cognitive challenge for input and output arguably decreased. This finding infers that the level of cognitive challenge within the thinking tasks is not long lasting. When this situation occurs, it is suggested that the *difficulty level* of HOTS can then be increased by intensifying the workload in the task. On the other hand, it could be sensible to move the topic from the concrete experience or value-laden to a more abstract concept. This management of the challenge level of tasks for optimum learning of the L2 language and HOTS as therefore very important in implementing any pedagogical innovation such as this.

Evidence in the video data as well as that in the interview and questionnaires showed that opportunities to argue and reason created by the socio-cognitive conflict explicitly in the thinking tasks enhanced students' argument-based reasoning skills. McGregor (2007) states these skills include rationalizing opinions, inferring, making deductions, making informed judgments or decisions and using precise language to reason. This result in the present study reflects the theoretical conceptual framework where students were expected to interact and argue with each other in group work when holding different values. As Kuhn (2005) stresses cognitive development is a benefit of

this type of argument. The present study supports Mercer et al. (1999) who found that children's individual reasoning improves in L1 classrooms when they are taught how to discuss and exchange ideas in groups. It also indicates that the HOTS approach can enhance students' argument-based reasoning skills in the L2 classroom as well as L1 classrooms.

The interview data provided some evidence of students' ability to apply these argument-based reasoning skills learned in this innovation to other domains. The result is encouraging although the number of interviewees reporting this was limited: the reliability of this finding needs to be considered. Though HOTS transfer has not been explored in L2 classrooms, several empirical studies completed in L1 classrooms reported transfer of reasoning and argumentation skills, for instance, Hunter-Grudin's (1985), Zohar's (1996) and Zohar and Nemet's (2002) work. In addition, Miri et al. (2007) found that with the promotion of HOTS in science class, students were capable of transferring these high cognitive thinking skills across domains. This study does triangulate with other researchers, for example, Watson (2000) and Billing (2007), who argue that transfer is promoted when learning takes place in a social context through collaborative methods. This finding implies that the HOTS approach can provide greater opportunities to apply argument-based reasoning skills learned to other fields.

Learning transfer is an ultimate goal when teaching students to think (Zohar and Dori 2003; Ritchhart and Perkins 2005; Reece 2007) and is a primary aim of HE in Taiwan. As Marini and Genereux (1995) argue, the education system is deemed to have failed when students cannot perform tasks outside of school from those learned in class. Without transfer, the work of trying to foster thinking might be in vain. Yet, the roads to transfer within and across domains are rather rocky (Ennis 1989; Salomon and Perkins 1989; Detterman 1993 cited James 2006). Inspiringly, this study showed some opportunities to allow the transfer of HOTS. Thus, future investigation into the extent to

which the HOTS approach enhances the transfer of argument-based reasoning skills from L2 classrooms to other domains is recommended.

Overall, the findings indicate the HOTS approach not only can enhance high cognitive thinking, but also provides greater opportunities for the development of problem solving, accelerated thinking, argument-based reasoning and the transfer of the reasoning skills. In addition, the results indicate that the cognitive challenge of the tasks can decrease after students become accustomed to and improve high cognitive thinking. The development of HOTS and the decreasing cognitive challenge are highly related to the progress made in speaking. L2 speaking can be seen as a cognitive tool. The sharper the tool is the better chance of advancing high cognitive thinking students gain. For that reason, it is now important to explore students' L2 speaking development.

5.4.2 L2 speaking development

The previous section showed how the contexts for socio-cognitive conflict created opportunities for high cognitive peer interaction. It is now necessary to examine whether these opportunities were conducive to L2 speaking and therefore the impact the HOTS approach had on speaking development. This will be explored by looking at the length of utterance, grammatical complexity and the overall speaking proficiency of students in the innovation and comparison classes. Language learning in L2 classrooms is with no doubt the most important issue in this study and without achieving a positive impact the effort L2 learners devoted to cognitive improvement will be in vain.

The video data showed great improvement in the innovation class than in the comparison in terms of the length of utterance. This was backed up by the results gained from the interview and questionnaire. This finding is in line with other research,

for example, Cole and Williams (1973) and Smith (1978) in L1 classrooms and Brock (1986), Nunan (1990), Alcon (1993) and Godfrey (2001) in L2 classrooms, where the length of utterance becomes longer after the innovation. It indicates that the HOTS approach provided plenty of opportunities for students to talk and try out their hypotheses about comprehensibility or linguistic structure, and the possibility to improve speaking proficiency is increased. With this approach, whether the more students elaborated, the better the quality of talk was and higher L2 speaking proficiency achieved will be explored in the following discussion.

It should be noted however that the length of utterance in the comparison class also increased during the period of study. It is not certain whether this improvement resulted from their normal learning process or thinking tasks during data collection phases. It is possible that the comparison class benefited from this relatively short exposure to thinking tasks. If this is the case, it would indicate that in situations where no modelling is carried out, the thinking tasks carry an immediate function of enhancing the amount of talk. It would be worth conducting a further study to clarify whether thinking tasks have such an immediate impact on encouraging students to talk more. However, the additional impacts noted in the innovation class would indicate a more fundamental change in the students' competence than just increased utterance: it is the quality of the utterance which is important.

The innovation class saw greater linguistic improvement with regard to grammatical complexity than the comparison class. This was also found in previous studies using higher-order questions, for example, Cole and Williams (1973) in L1 contexts, and Brock (1986), Nunan (1987, 1990) and Godfrey (2001) in L2 classrooms. This improvement could be due to the need for students to explain ideas more completely and logically as part of the thinking tasks, meaning students used more complicated syntactic structures. Through practicing during the innovation, students had more

opportunities to apply the language rules appropriately and spontaneously in different contexts. Thus, more syntactically complex responses were triggered. This fits with Coyle's (2002) assertion that unless grammar learning is contextualized through meaningful language use and learners take ownership of the language, the grammar per se remains a conceptually difficult irrelevance. The finding infers that the HOTS approach encourages students to utilize more complex syntax and therefore gives some indication of impact on quality and as well as quantity in the innovation class.

The above discussion shows useful and corroborated evidence to support the idea that the HOTS approach is effective in enhancing students' talk in terms of the linguistic competence, grammatical complexity. Yet, it needs to be noted that this improvement alone does not necessarily represent a good command of spoken language. As Ortega (2003) found syntactic complexity does not necessarily have a positive relationship with L2 proficiency. Therefore, it is now important to explore the speaking proficiency as a measure of the extent to which students have good command of English, including four aspects namely fluency and coherence, lexical resource, grammatical range and accuracy, and pronunciation. This will present a better picture of students' speaking ability.

The current study found that the innovation class outperformed the comparison class in terms of overall speaking proficiency when data from the video were analysed. This improvement is relevant to the theory of second language acquisition output hypothesis. Swain (1985) argued that the need to actually produce language forces learners to think about syntax, thus enhancing L2. In Swain's (2000, 2005, 2006) later work, she considers second language acquisition from a sociocultural theoretical perspective and suggests that when spoken language serves as an intellectual tool, it assists L2 learning process. The target language used in this study completed and transformed thoughts (Vygotsky 1987), and it shaped and reshaped cognition. In such

cognitive activity numerous speaking opportunities were created for students to try out their linguistic hypotheses.

Students in the present study listened to the speaker's thoughts, and thereby they had opportunities to notice features of the target language needed. Gass (1997) states these stages of the process learners pass through for second language acquisition as the move from noticing L2 input to producing L2 output. As students consider each linguistic unit, they then attempt to reproduce it in a new context when expressing their point of view. Alternatively as the speaker rephrased, repeated and reorganized linguistic units to make their thoughts more comprehensible and logical to others; they were seen to perform a syntactic and semantic analysis of the language. This cycle, from noticing L2 input to output, allowed students to learn from their peers and try out their language hypotheses in the space created for enhancing speaking by the thinking tasks. This finding supports Lin and Mackay's (2004) and Paren's (2003) claim, that infusing thinking skills into L2 classrooms not only gives learners the space to think out loud, but also adds linguistic value to the classroom.

The thinking tasks arguably helped students to move away from transactional language and rote learning towards meaningful interaction enabling students to engage with the language at more than a superficial level. Exploration of students' overall speaking proficiency in the field of using higher-order questions in L2 classrooms has not been found in the literature. Thinking skills programs conducted in L1 classrooms have shown positive results on student oracy performance. For instance, Bowdler et al. (1992) found students became better at elaborating their reasons and justifying their opinions, and Lipman et al. (1980) discovered that student oracy could be improved through 'Philosophy for Children' when students became accustomed to elaborating and asking for reasons and opinions, and were asked to build on each others' ideas. This study therefore finds that the HOTS approach like

other thinking skills programs having the impact on L1 development enhances L2 speaking proficiency.

The real issue for studies of second language learning is the extent to which a treatment truly resulted in learning by exploring whether the effects are long lasting, rather than simply looking at the immediate effect, as Mackey and Gass (2005) argue. Therefore, it is gratifying to note that the video data showed that students' speaking proficiency was maintained after the innovation finished. Investigation of published empirical studies in the field has revealed no research examining this kind of effect in L2 classrooms. However, this finding is in agreement with the claim made by Sousa (2001) of L1 classrooms: the use of HOTS increases understanding and retention. The thinking tasks required students to call upon existing knowledge, concepts, skills and strategies in terms of both language and content and according to Met (1998) it is such strong connections amongst concepts and knowledge that increases learning and retention. This result indicates that the HOTS approach not only has an immediate effect on L2 speaking, but the effectiveness for L2 learners is long lasting. Such sustainability is very important for a successful teaching approach.

Compared with other tasks conducted in L2 classrooms to improve L2 speaking, the thinking tasks arguably have greater benefit for students. This can be explored by examining effects carried out by different tasks. The literature shows that different tasks enhance different perspectives of L2 speaking. Evidence in the current study showed the thinking tasks not only allowed students to explore the language needed for structuring and expressing their thinking, but also provided many opportunities for complex syntactic processing for language development. Yuan and Ellis (2003) investigated the effect of pre-task and online tasks on language production and found a variation in impact. The pre-task planning enhanced grammatical complexity while online planning influences accuracy and grammatical complexity. Similarly decision-

making tasks have been found to have greater effect on language complexity than personal and narrative tasks (Foster and Skehan 1996; Skehan and Foster 1997). In addition, Bygate (1999) has examined the impact of unscripted narrative tasks and argument tasks on language development in secondary school and found that the former seemed to stretch the speakers more in terms of syntactic complexity and lexical processing, while the latter have a premium on brevity and lack of elaboration, which appears to push students towards less complex syntactic processing. These studies all found tasks which were efficient at enhancing certain skills, e.g. syntactic complexity and accuracy. However, this research shows that thinking tasks mean learners are stretched to use a wider range of vocabulary items and increase the amount of talk, grammatical complexity and the overall speaking proficiency. The HOTS approach achieved a much more comprehensive improvement than other speaking tasks.

Impact across the ability spectrum

The findings obtained from the video data also showed that speaking improvement was generic; that is, the majority of students in the class improved. Firstly, this means that no student in particular will dominate the interaction when conducting the HOTS approach. Everyone, including students with higher and lower speaking proficiency had equal chance to elaborate their thoughts. When they did use these opportunities, it was seen that each opinion was valued and justified regardless of individual proficiency differences, which could be attributed to the modelling in a teacher-led setting. As McGregor (2007) and Mercer (2000) argued, respecting each other's opinion is essential in the learning process when students are working together with thinking activities. Secondly, the finding fits with empirical studies, such as Sizmur and Osborne (1997), show that collaborative elaboration is positively related to group performance. In addition, it suggests that the collaborative learning that occurred as a

result of thinking tasks is positively related to individual learning outcomes. This is a further evidence that HOTS can develop individual L2 speaking competence.

Such individual L2 speaking enhancement can be triangulated with the significant gains by the lower achievers (p. 115) from video data. This confirms the literature reviewed in Section 2.2.1 which showed that socio-cognitive approaches facilitate the learning of low achievers. It also supports previous research, such as Quicke and Winter (1994) and Powell and Makin (1994). The use of social constructivist approaches in this study helped low achievers and they became more able to share ideas, express opinions, and talk about their learning. On the other hand, this finding is in contrast to Kowal and Swain's (1994, 1997) and Yule and Macdonald's (1990) research, which showed less proficient students engaged in little negotiation with their peers: either they were willing to let the stronger students do the task or they were not allowed to be involved.

The findings also evidenced that the lower achievers made greater improvements than the higher speaking achievers. This result is in contrast to Watanabe and Swain's (2007) findings where both lower and higher achievers improved, yet higher achievers benefited more in their L2 learning. There are two possible reasons for this. One is that the groups in the current study consisted of three students and this allowed the lower achiever to gain more time to process the information and reflect on the language use while the other two higher achievers were talking. Another is that thinking tasks motivated students either higher or lower achievers to talk. As discussed in Section 5.2.1, wait-time is essential in the successful implementation of high cognitive thinking (Tobin 1987) and Sousa (2001) also argues that when a lower ability learner is able to process a greater amount of information at a certain thinking level, their thinking and achievement can be enhanced. Overall, this finding is consistent with those researchers (for example, Watanabe and Swain 2007; Shima 2008) who found that

proficiency differences do not necessarily affect the nature of peer assistance and L2 learning; students at all levels can be benefited.

Despite the listening problems encountered by certain low achievers (Section 5.4.3), these changes in the learning process suggest that the HOTS approach is not the sole prerogative of more capable learners, but should be expected at all levels. Advantages for both higher and lower achievers can be roughly equally distributed with the HOTS approach. This idea is supported by Pogrow (1996) and White and Frederiksen (1998) and is exemplified by Zohar and Dori's (2003) claim that HOTS training is good for all levels of learners.

In addition to the issue of language development among low achievers, empirical research has also shown that in L1 classrooms the effect of HOTS on lower achievers' cognitive level and academic performance is positive (Zohar and Dori 2003; Pogrow 1996). Vygotsky (1978) and Watson (2000, 2001) also believe that a social constructivist approach could promote low achievers' cognitive development. Nevertheless, the current study did not further look into this aspect. This would certainly be worthy of future study. Most importantly, the HOTS approach may have the valuable effect of reducing the educational disadvantage of low achievers while also being beneficial for high achievers. Further exploration of the effect on low achievers' L2 performance would be useful.

The above discussion revealed the impact of the HOTS approach on L2 speaking proficiency not only related to the group but also individual performance and had long lasting effects. This positive finding suggests that it is worth introducing HOTS into L2 classrooms in order to expand the learning dynamic and go beyond simply learning the target language itself. Van Duzer and Florex (1999) claim that, since the world has become overwhelmed by information, how people process it has become more

important than the information itself. It cannot be denied that the top priority in L2 classrooms is to enhance language proficiency. Thus, instead of merely using basic literacy skills, language teachers should encourage students to go beyond the surface and discover deeper meanings which require high cognitive thinking. In this study, in achieving the goal of L2 learning, students also developed their higher-order thinking. As a result, it is surely appropriate for L2 learners to receive such high cognitive training.

Relationship between interaction and L2 speaking

The thinking tasks offered speaking opportunities, yet it was high cognitive interactive talk that was generated through a combination of characteristics which revealed their potential and increased the level of cognitive challenge. Such learning contexts were seen to directly impact on students' speaking performance. Therefore, it is now worth examining whether the relationship between such interaction and speaking proficiency was positive. The video evidence demonstrated that when high cognitive peer interaction increased, L2 speaking proficiency, grammatical complexity and length of utterance were also enhanced. Associated with this finding empirical studies (for example, Loschky 1994; Mackey 1999; Sato 1988) have shown that not all communication among learners can be said to facilitate linguistic development. It depends on the specific interaction pattern that occurs. Results found by Zhou (1991), Gass and Varonis (1994), Seliger (1977) and Mackey (1999) show peer interaction can have a positive relationship with L2 achievement which corroborates the finding of this research.

Several empirical studies of L2 learning have supported collaborative interaction between language learners assistance can be provided equally between non-experts (e.g., Donato 1994; Ohta 1995; Swain and Lapkin 1998). Thus, high cognitive peer interaction with peer scaffolding, as Krupa-Kwiatkowski (1998) and Donato (1994)

argue, could trigger students' cognitive processes thus providing an interaction pattern that is more conducive to language learning. This study shows that the high cognitive peer interaction pattern created by the HOTS approach appears to facilitate L2 linguistic development and speaking proficiency.

It should be pointed out that interaction utilizing low cognition may also contribute to the improvement of L2 speaking. The video data showed that when students encountered problems, for example different interpretations of a word or unclear message, low cognition was applied for interaction. This is relevant to the interaction hypothesis, a theory of second language acquisition. In formulating the interaction hypothesis Long (1981, 1983, 1985, 1996) argues that comprehension is made possible, and is even facilitated, when interactional adjustments are present. In this study, the HOTS approach directly involved learners in the process of negotiating meaning; where input was made comprehensible as a result of modification. When communication problems arose, such as a lack of clarity in the message, unclear pronunciation and the problem of word interpretation, students negotiated and clarified using low cognitive thinking. It provides the necessary and sufficient conditions for acquisition and mastery of a second language. Long's hypothesis clearly stated negotiation of meaning, the type of interaction which requires lower-order thinking, enhances L2 learning. Confirming Long's finding, this study also found that it was the high cognitive interactive talk, the type of interaction which requires higher-order thinking, that creates numerous speaking opportunities and these output opportunities, as argued by Swain (1985), allow students to try out their linguistic hypothesis and facilitates L2 learning. This indicates that peer interactions including high cognitive and low cognitive have their respective own roles in contributing to L2 speaking enhancement. Peer interaction with low cognition makes the input comprehensible, while high cognitive peer interaction provides students maximum opportunities to talk and try out their hypotheses about comprehensibility and linguistic structures.

Maturation of speaking proficiency

An exploration of how students in the innovation class increased their speaking proficiency through interaction showed that at the start of the innovation, students' utterances were less coherent with few completed sentences and very little 'real' dialogue. Since coherence is a sign of development in communication and comprehension (Schiffrin 1994), this indicates that communication at this initial stage was not mature. At post-test the coherence was more developed, including both the fluency and clarity of argument. Students did not just focus on using correct sentences, grammar and vocabulary, but on ensuring that what they said made sense. This as what Allwright (1979) contends is meaningful interaction that helps to enhance L2 speaking. At the same time, students' high cognitive interaction increased. Speaking language used in this innovation served as a cognitive tool and L2 knowledge was co-constructed and developed by students interacting with their peers and reflecting on their own language use (Swain 2000, 2005, 2006; Brown and Yule 1983). The high cognitive peer interaction pattern that occurred in this innovation shaped the talk in group discussion and most importantly, facilitated the enhancement of L2 speaking proficiency.

In contrast, speaking proficiency decreased in the comparison class. This could be due to the lack of negotiated meaning. In the comparison class, students' pronunciation at pre-test was good and sounded more English than post-test. The students also seemed to have more confidence and used more vocabulary and utterances were much shorter at this stage, hence the impression that pronunciation was clearer. At post-test, however, pronunciation was seen to be poorer and although students elaborated their thinking more, they did not seem aware of whether their pronunciation or the meaning lacked the clarity to get their ideas across. It can be concluded that the interaction pattern of high cognitive monologue students has experienced meant speakers had not received any feedback on ideas or pronunciation

from the interlocutors, for example, 'can you say it again?' or 'what do you mean by..?'. Without this feedback, the speaker would not be able to realize that his/her own pronunciation could cause problems for listeners or that the meaning delivered was not sufficiently clear.

In short, the high cognitive peer interactive talk that occurred in the innovation class can be seen to have a positive relationship not only with cognitive development, but also with L2 speaking proficiency. This successful result can be attributed to high cognitive peer interaction where numerous opportunities to try out language hypothesis were provided and peer interaction with low cognition where input were made comprehensible. It thus can be concluded that the HOTS approach has the attributes to enable learners to communicate with high cognition and low cognition and such interactions contributing greatly to students' L2 achievement.

L2 transfer

The authentic scenarios created by thinking tasks impacted on L2 transfer. The findings obtained from the interview data showed students perceived that language transfer occurred. Two participants reported they were able to apply the target language learned to other fields, for example online communication. James (2006) states that the notion of *authenticity* is fundamental in this type of learning. When it matches with instructional materials and tasks in L2 classrooms, then students are more likely to see the relationship with the world outside the classroom and this may stimulate learning transfer. This finding can be triangulated with the literature regarding the relationship between pedagogy and language transfer. Jones et al. (1987) believe that a pedagogy mixing thinking skills and language learning promotes language transfer. This evidence suggests that L2 transfer could occur through the use of the HOTS approach. However, the reliability of this finding needs to be considered due to the limited cases reported. Overall, this finding indicates that the HOTS approach may

provide a greater possibility for L2 transfer, but further research is essential within the L2 context to explore the generalisability of this phenomenon. James (2006) has reviewed some techniques for maximizing the potential of teaching for transfer in general education and this could be referred to when designing instruction to promote this transfer of learning.

5.4.3 Learning behaviour

This section will firstly discuss the extent to which the HOTS approach influenced students' learning behaviour: their disposition towards taking risks and awareness of the process of learning. The impact on students with low listening comprehension is then explored.

Evidence from the interview data revealed that students' disposition towards risk taking was developed. The thinking tasks were designed to require students to articulate their own thoughts and reach a consensus. When the desire to express thoughts occurs, it stimulates students to talk. In the present study students overcame a psychological barrier in terms of speaking in the target language and intellectual barriers when they started to think out loud and take on board the challenge involved. The change in observed behaviour, breaking through these barriers, as found in Casanave (2010), enhanced risk taking. Arguably, the environment created by the use of group work produced less anxiety and facilitated this action. Chen (2002) observed that Taiwanese students can fear oral activities in L2 classrooms, a finding corroborated by the exploratory study (Section 1.3), indicating that students lacked risk taking ability. Similarly, participants in the current study claimed that at the beginning of the innovation they were afraid of making mistakes and did not have the courage to talk. It is through the HOTS approach training and the nature of thinking tasks they started to give it a try. This can be seen as a milestone in the learning process of L2 speaking, where students move from 'no courage to talk' to 'not being afraid of

speaking out and make mistakes'. This result implies that the HOTS approach encouraged students to change, take risks and step into the real realm of L2 speaking, even though the utterances produced can be segmented.

The frequent use of high cognition could have impacted on students' learning awareness. The interview and questionnaire data showed that the majority of students in the current study were aware of their own learning. One case found that students in the comparison class perceived their improvement in answering higher-order questions to be twice as high as the innovation class, which was in contrast to the results obtained from the video data. This anomaly could be explained because the comparison class did not receive the training of the HOTS approach and therefore did not experience additional thinking tasks, outside those used for data collection and very often students could find answers to the questions posed in the textbook. Thus, they did not encounter as many difficulties as students in the innovation class. Consequently, they felt more satisfied with their performance. In contrast, students in the innovation class who had encountered more difficulties because of the authentic scenarios created by the thinking tasks, recognized that there were still a lot of space for improvement and their expectation became higher. Thus, they evaluated their improvement considerably lower. This finding is supported by Watson (2000), Leat (1998), Lin and Mackay (2004) and Moseley et al. (2004) who have all showed that thinking skills activities used in association with a social constructivist approach can assist learners in becoming more aware of their learning. This implies that frequent exposure to the HOTS approach can make students become more aware of what they cannot do in the L2 classroom, especially in terms of linguistic capacity and expressing thoughts more completely and logically in the target language.

In addition to the influence learning contexts had on students' learning behaviour, the finding obtained from the interview data showed there was an additional impact on

students with low listening comprehension. Two students reported that they were not able to understand the questions or to follow the lesson in teacher-led settings. As Godfrey (2001) found, students' English proficiency can hinder students from participating fully and also limit the degree to which they benefit from higher-order questions. This appears to be the case here. Therefore, any teacher considering a HOTS approach needs to take this issue into account. This problem might be solved by employing questioning strategies (Wu 1993), such as rephrasing, simplification, repetition and decomposition, both by the teacher or students. With further explanation, these students will still have the chance to elaborate their opinions. In addition, providing more in the way of related sentence patterns and vocabulary items might help students elaborate their ideas.

In general, the HOTS approach enhanced the students' dispositions towards taking risks and developed students' learning awareness. Nevertheless, any impact on students with low listening comprehension cannot be disregarded and it needs to be taken into account when encouraging teachers to use the HOTS approach. So far, this study has shown that the HOTS approach is conducive for both L2 speaking and thinking and positively impacted on students' learning behaviours. Yet, the practicability of applying thinking skills into an L2 classroom also depends on students' attitudes towards it. Thus, next section will discuss students' attitudes towards and perceptions of the HOTS approach.

5.4.4 Attitudes and perceptions

The design and implementation of thinking tasks, the learning contexts created and the achievement students gained could all be seen to affect students' attitudes towards and perception of the HOTS approach. The finding obtained from the interview and questionnaire data showed that the majority of students held positive attitudes towards the HOTS approach: 72.5% of the students found it interesting and

approximately 60% of those surveyed had the experience of enjoyment in answering higher-order questions in English. It implies that students were able to adapt to high cognitive training in the L2 classroom. This finding is in contrast to the results of Tan (2007) who investigated students' attitudes towards higher-order questions and found that students resisted. Students in the current study came to like this approach and articulated their thoughts increasingly in English over time, as shown in video data. This could be due to the gain they perceived or benefit they saw in their own and others' ability. In fact, Green (1993), investigating how much learners' enjoyment of communicative and non-communicative activities corresponded with their perceptions of the usefulness of these activities, found there was a general tendency for effectiveness and enjoyment to be highly correlated. Indeed, apart from the enjoyment students had, the video data also shows a positive effect on students' performance.

On the other hand, comments on the 'interest' level of the tasks were in line with Lin's (2005) study where students found it interesting because they could talk about their thoughts towards movies shown in the class. One possible reason to explain why students in the current research found it interesting and were more motivated towards English learning could be the variety of tasks and materials employed. The teaching method described in Tan's (2007) study was related to a teacher merely probing higher-order questions. Any activity like questioning, as Kissock and Iyortsuun (1982) observed, used all or a large portion of the time becomes unproductive. Unlike Tan's (2007) study, the current research and Lin's (2005) study both used a range of material and a variety of tasks to encourage higher-order thinking. Thus, students arguably did not feel fatigue and therefore a negative attitude.

The HOTS approach however was not universally popular. The interview data revealed that one-quarter of the students resisted this approach and preferred the teaching method which was similar to grammar-translation. This result is consistent

with Tan's (2007) finding where students resisted higher-order questions; they did not like to articulate how they reached the answer to the question. In the current study students provided three reasons for this resistance. Firstly, they felt this teaching approach was challenging and required more effort. Secondly, they did not consider that they learned any more than using the old methods. This could be that their learning styles enabled them to benefit more from the more didactic teaching approach similar to grammar-translation (Coffield et al. 2004). A third reason given was that the 12-week innovation was not long enough for these students to adapt to or fully benefit from it, and this could be corroborated by the work of Eisenman and Payne (1997).

The problem of a negative attitude could be solved by two means. One is to include other types of thinking tasks (as discussed in Section 1.5.1), such as problem solving task so as to increase the variety and the likelihood of a match with students' preferences. The other possible solution could be offering a mixture of pedagogies. It can be argued that there is no 'one size fits all' pedagogy. As Nakatani (2005) stresses, training does not always improve learning for all students. Thus, to meet the demand of different learning styles, applying different types of pedagogy can be considered.

Having explored students' attitudes towards the HOTS approach, it is now appropriate to discuss their perceptions of higher-order thinking skills used in the L2 classroom. It is interesting to note that results from the ANCOVA test on the questionnaire data showed a significant difference between the two classes (p. 151). Students in the comparison class considered the importance of HOTS much higher than those in the innovation class. A possible explanation for this could be that students in the comparison class had experienced the advantage of thinking tasks during data collection phases and compared with the tasks in the textbook, they may find this

approach more exciting and motivating with its greater opportunities to talk. This was confirmed somewhat by the video data.

Although the innovation class receiving HOTS training did not evaluate high cognitive thinking as more important after the innovation, this does not mean they disregarded or did not perceive the importance of it. Rather, they saw the significance through realizing the lesser importance of lower cognition. As Paired-Samples T Test indicated that there was an impact on the innovation class at post test (p. 153): the students did not consider lower-order thinking skills to be as important as before. It might be due to the fact that the students had previously had strong beliefs about reciting vocabulary and understanding texts when learning a foreign language. After the innovation they appreciated the positive effect of using HOTS. These findings indicate that the lower importance attached to low cognitive thinking reflects the importance of high cognitive thinking. This backs up the finding where students had positive attitudes towards the HOTS approach.

It should be noted however that in L2 learning, some commentators argue that lower-order thinking skills are the fundamental skills required. For example, Ding (2007) explored the learning strategies used in L2 learning by successful learners and found that memorization and imitation were the fundamental skills used for acquiring linguistic competence. In addition, Bloom (1956) has stated, learners need to use lower-order thinking skills to recall from their memory and to activate their schemata, so high cognitive skills can be operated. Overall, positive attitudes towards the HOTS approach indicate that students can adapt to this high cognitive training and find it useful, although an attitude of resistance also existed. There is no doubt that the range in attitudes need to be taken into account when implementing this approach.

5.4.5 Reflecting on L2 speaking problems in Taiwan

Having explored the design and implementation of thinking tasks, the contexts created, the learning opportunities, the positive outcomes gained and students' attitudes and perceptions explored, it is now important to appraise whether the HOTS approach can provide solutions to L2 speaking problems observed in Taiwan (Section 1.2 and 1.3).

The HOTS approach was particularly designed for L2 speaking classrooms. It can be conducted in a class or small group discussion forum. The former provides opportunities for students to interact with the teacher, while the latter provides individual student with the increased chance to use the English language they are learning. Topics used in the current study were related to life, for example travelling experience, the ideal mate, characters, travelling packages, pets, etc (Section 1.5.1). Such topics enabled students to activate their schemata so that they could come up with their thoughts. This led to the authentic contexts being created which were similar to everyday life (Section 5.2.2). As students claimed in the interview, the discussion was like having a chat with friends in a natural setting. Opportunities for talk were provided not only by the thinking task itself, but also by high cognitive peer interaction (Section 5.2.2). Students were required to think hard, so they could generate new ideas and verbalize them. Such learning contexts created addressed the issues of insufficient speaking opportunities (result found in the exploratory study), large class size (Ho 2002; Xu 2002), material unrelating to students' experience (Shou 1995; Yang 1992) and lacking daily life context (Shou 1995).

The problems regarding students' learning behavior namely low motivation (Moslehpour and Chou 2004; Ho 2002; Xu 2002; Shou 1995; Yang 1992), high anxiety (result found in the exploratory study), low risk taking ability (result found in the exploratory study) and passive learning attitudes (Shou 1995; Yang 1992) could be

tackled by the HOTS approach. Group discussion created an environment with less anxiety (Section 5.2.1). The tasks focused on the expression of meaning, so some of the students' anxiety related to the need to be grammatically accurate may have been lessened. Grammar errors were corrected after the discussion finished, in a teacher-led setting, where students felt much less anxious. The design of the thinking tasks, requiring students to talk about their thoughts encouraged students to take risks in which they started to think hard and try to speak out their thoughts (Section 5.4.3). Their motivation to talk and confidence to interact were increased (Section 5.2.2 and 5.2.3) as they found the target language being used authentically and the tasks more interesting and enjoyable. The learning contexts created by the thinking tasks were learner-centred where students could dominate the talk, initiate and control the topic. In this learner-centred context, students were thinking proactively (Section 5.4.1). This arguably encourages active learning.

Finally, the issue of low speaking proficiency (Nunan 2003) was tackled by the learning outcome gained in this study (Secgtion 5.4.2). The HOTS approach proved to be considerably effective in enhancing L2 speaking. Students were able to argue and reason in which they had a good command of spoken language and were able to elaborate more with a higher syntactic complexity. The HOTS approach provided some solutions to L2 speaking problems and the space for utilizing high cognitive thinking in Taiwan.

5.5 Summary

This chapter has answered the research questions by discussing three themes: contexts created for socio-cognitive conflict, high cognitive interactive talk and performance outcomes. Encouragingly, it shows that infusing thinking skills into a Taiwanese L2 classroom is useful and practical. The HOTS approach was effective in terms of enhancing students' speaking and thinking performance with the majority of

students having positive attitudes towards this approach. In contrast to Atkinson's (1997) claim, the result of the present study supports Day's (2002) finding where Asian students can be trained to think critically. As Littlewood (2000) and Stapleton (2002) found that Asian students do not wish to be spoon-fed with facts but would like to be active and independent in learning. What they need is the opportunity to learn to think and learn to use their high cognition and express themselves clearly in the target language and such opportunities can be provided by the HOTS approach. The Taiwanese educational goal to reduce the use of rote learning and equip students with thinking skills is very likely to be fulfilled by this approach.

The thinking tasks provided an optimal L2 learning condition where the need to use the target language was met and opportunities to initiate and control topics were provided. Students had opportunities to think critically and explore authentic situations through dialogue. They were also encouraged to change and take risks and this arguably enabled those having fear of English speaking to start to talk and enter the real realm of L2 speaking. Their confidence in speaking in the L2 was increased and thus participating in peer interaction was also encouraged.

One significant result was the occurrence of socio-cognitive conflict. The thinking task itself offered opportunities to talk and think; yet it was high cognitive interactive talk emerging from socio-cognitive conflicts further increased the opportunities. There were two types of peer interaction occurred in this study: interaction with high cognition and interaction with low cognition. Peer interaction with low cognition provides comprehensive input while that with high cognition, very importantly, gives students plentiful chances to try out their linguistic hypotheses and think critically. Collaborative learning including peer interaction with high cognition and low cognition that occurred as a result of the thinking tasks positively relate to both thinking and speaking performance. The impact on speaking performance not only associates with group

performance but also individual learning outcomes. It is important to note that peer interaction with high cognition was attributed to the modelling in a teacher-led setting; without the modelling such interaction might not occur.

High cognitive interactive talk enhanced students' linguistic competence and enabled them to have a good command of spoken language; their length of utterance became much longer and contained a much higher complexity of syntax with much better coherence. Very importantly, such effects were long lasting and generic. The approach benefited not only high achievers, but also low achievers, and this implies that the HOTS approach is expected to be useful at all levels; it may have the valuable effect of reducing the educational disadvantage of low achievers while also being beneficial for high achievers. Besides, greater opportunities for enhancing students' L2 transfer were provided due to the authentic scenarios created. Further, when compared with the effectiveness of other speaking tasks undertaken in L2 classrooms, this study achieved a much more comprehensive improvement.

Apart from the dividend yielded for L2 speaking, students also made significant progress in cognitive performance. They produced a lot more HOTS and very significantly, they were able to use these high cognitive skills for interaction in a social context. This trained students to think independently, proactively and simultaneously in social contexts and such ability could be maintained after the innovation finished. In addition, their thinking seemed to become faster and skills like problem solving and argument-based reasoning were strengthened. With these positive outcomes, it is not surprising that the majority of students (around three quarters) responded positive attitudes towards this approach and found it interesting.

Inspiringly, the HOTS approach could provide solutions to L2 speaking problems: issues regarding to learning contexts, e.g. a need of contexts similar to real life,

learning behaviours, e.g. low risk taking ability, and learning outcome, e.g. low speaking proficiency. An optimal L2 learning context allowed students to communicate and use the target language more authentically. Students could dominate the talk and initiate topics of interest. Besides, their confidence in L2 speaking and their disposition to taking risks were enhanced. The nature of the thinking tasks encouraged active learning. That L2 speaking enhanced greatly addresses the issue of low speaking proficiency faced by the majority of Taiwanese students.

Overall, the HOTS approach had strong impact on students' L2 speaking and cognitive performance with students' holding positive attitudes. Very importantly, it also can address L2 speaking problems in Taiwan and tackling current educational goals. The related implications of this study and the knowledge contributed to the field of using high cognition in L2 learning will be further discussed in Chapter 6.

CHAPTER 6 - CONCLUSION

6.1 Introduction

This study investigated the practicability of infusing thinking skills into a Taiwanese university L2 classroom through an innovation of the HOTS approach. A mixed-method approach with a case study design was used to investigate the effect on students' speaking and thinking and their attitudes. The results reveal that it is workable to infuse thinking skills into L2 learning to enhance students' speaking and thinking performance. The findings of this study also raised a number of issues, which led to the following recommendations and suggestions for future research of this study.

This chapter comprises the concluding remarks of this study. The limitations of this study will inform researchers what can be done to achieve a more rigorous and comprehensive research design in the future. The implications of the study and recommendations are highlighted to notify policy makers and teachers. Areas arising from this study which may provide a basis for further research are also provided for researchers in this specific field, with the final section dedicated to the knowledge contributed to the field of using high cognition in L2 learning.

6.2 Limitations of the study

This study was designed to acquire a greater understanding of introducing thinking skills into the L2 classroom. The findings from this study arguably accurately reflect the effects of the HOTS approach on students' performance and participants' attitudes, but there may be limitations due to the narrow range of data collected and the time constraints of this research. The limitations of this study are stated as follows.

The first limitation is the length of innovation conducted. Eisenman and Payne (1997) found that a HOTS program is more effective for young adolescents after two years of treatment. Due to the time constraint, this innovation was for just 12 weeks. This thinking skills approach is different from students' previous learning method, which tended to be rote learning. It might require a long time for some students to adapt to active learning, from rote learning, although some impact has been found regardless of the relative time constraints.

Second, a potential weakness in the design of this study is lack of standardization between the study group members in the comparison and innovation class prior to the commencement of the research. Due to the limited number of students volunteering to participate as study group members, a selection of students with equivalent English speaking proficiency was not conducted, although the average scores of the English language in the two study groups were approximately the same in the first semester. To reach a higher validity of a study, it is necessary to have study group members with equivalent speaking ability prior to the commencement of the study.

Third, the original research design was to collect 15 minutes of data but because of the technical problem occurred during recording (Section 3.7.1), some groups were only recorded for eight minutes. In order to maintain the consistency of the recording length in both pre- and post-test, the data used for analysis were kept at eight minutes long. The significance of the effect gained may be different if 15 minutes data had been available to be analysed.

Participants recruited in the present study majored in different subjects; the comparison class majored in Maths, while the innovation class majored in Visual-Art Design. Their thinking styles can be different: Maths students are generally better at analysing figures while Visual-Art Design students are good at creating images, and

this may have affected the results gained. Selecting participants with the same major subject would have given a more exact appreciation of the effect of this approach.

In addition, a further limitation of this study revolves around the subjects. The theoretical generalizability of the results needs to be considered because the sample is exclusively non-English majors, which forms the majority of L2 learners in Taiwan universities. Replication using samples made up of English majors would extend the understanding of the effect of the HOTS approach in the university L2 setting.

Finally, with a small sample size caution must be applied when interpreting the theoretical generalizability of the HOTS approach in the university L2 setting. However, a comparison group was included therefore the reliability of this study was strengthened.

6.3 Implications for HE

Pedagogy for L2 speaking in Taiwanese HE needs a fundamental rethink and a real change. As reviewed in Section 1.3 and 1.4, the current teaching methods do not really provide optimal learning contexts for students to practice speaking. Students need opportunities for authentic language use and to be able to dominate the talk and converse as in a daily life setting. It is the HOTS approach which can provide such a real change to improve the learning.

Enhancing students' thinking ability is the goal of HE; nevertheless, how thinking skills can be infused into content teaching has not been instructed explicitly. It seems that there is a gap between the goal and the implementation. This study has revealed how to integrate thinking and L2 learning by using the HOTS approach. This approach arguably helps L2 learning and fulfils the goal of HE.

6.4 Recommendations for teachers

For thinking tasks to be effective, teachers must view teaching as a process of developing and enhancing students' ability to learn. The teacher's role is not to transmit knowledge, but to serve as a facilitator for learning (Lin and Mackay 2004). From the social constructivist view of learning, teachers are expected to provide scaffolds to support learning, encourage learners to discuss ideas and share understanding and expect learners to mediate each others' ZPD (McGregor 2007). Such kinds of teacher behaviour may stimulate students' active participation where students' thinking can be provoked and inspired. On the other hand, it is also important to explicitly explain the aim of using the HOTS approach and the possible effects on students so that the students are made more aware of HOTS being used in their learning process and the potential impact on their performance. This helps students understand the reasons for introducing thinking skills into L2 classrooms.

Regarding the implementation of thinking tasks, it could be difficult for teachers who are not familiar with the concept of HOTS to design thinking tasks and carry out the teaching. Nagappan (2001), in exploring the implementation of a HOTS program in Malaysian classrooms through a survey questionnaire, found that teachers experienced difficulty in acquiring the pedagogical knowledge to teach higher-order thinking skills, although support had been provided to the teachers. In addition, the act of choosing the innovator (Section 1.5.3) also indicates that teachers might need more time to become familiar with the teaching of thinking. These problems may be resolved with proper teacher training programmes, for example through conducting workshops, explaining the rationale of designing thinking tasks for L2 classrooms and offering direct experience with designing and teaching the tasks. This would, however, need support from curriculum designers, be they in the government or the universities themselves, to carry out programmes of professional development of teachers.

To successfully train teachers in how to implement the HOTS approach in L2 classrooms, Adey's (2004) model of effective professional development is worthy of consideration as it points out relevant factors influencing the implementation of an innovation. The present study has highlighted the importance of teacher behaviour and material design for enhancing student achievement. However, a successful professional development programme, aiming for positive student achievement, requires not only a change in teacher behaviour but also larger contextual changes, for example, the department and the university, as identified by Adey (2004). Nagappan (2001) found that despite of support provided, teachers find it difficult to adopt critical thinking into their content areas. It suggests that schools' commitment and support is not sufficient to lead to a success of professional development programme, if teacher behaviour does not change. This resonates with Adey's (2004) conclusion that any factor in association with professional development (such as teacher behaviour, collegiality and school's commitment) operates negatively, there will be little or no effect on teachers, and therefore on students. Thus, relevant factors need to be taken into account while conducting professional development programmes. In addition, teachers need to be aware of the wider repercussions of a HOTS approach. As indicated earlier (p. 71), the relationship between the teacher and the students was altered, with a shift of power and responsibility to the students. Such changes cannot be completely isolated and one must expect some long-term consequential impact which alters the identity and status of teachers.

6.5 Suggestions for future research

In conducting the HOTS approach, areas needing future research have been discussed in Chapter 5 and they are:

a) applying this approach to L1 classrooms to enhance student oracy, cognitive performance and academic achievement because cultivating students' thinking skills is an important issue in HE;

b) examining the impact on individual cognitive performance which enhances our understanding towards the cognitive development in L2 classrooms;

c) investigating the extent to which this approach enables L2 transfer which is important because students need to be able to use the target language outside the classroom through the use of this approach or their effort devoted will be in vain;

d) examining the generalisability of L2 enhancement, this enables researchers of the field a wider understanding towards the effect of the HOTS approach;

e) probing the effect on lower achievers' L2 and cognitive performance which allows a greater understanding towards the effect of the HOTS approach towards on lower achievers;

f) clarifying whether the thinking tasks have an immediate impact on encouraging students to talk more in situations where no modelling is conducted, this enhances our understanding towards the weight of thinking tasks and modelling in terms of the effectiveness; and

g) examining the effect of 'Odd One Out' task on L2 learning because this task works well in L1 classrooms and it is necessary to figure out its effect and the reasons why it worked less well in L2 learning.

With regard to methodology for future research, limitations of this study (Section 6.2) have informed what can be enhanced to achieve a more rigorous and comprehensive research design. There are additional issues to be considered: the length of innovation and analysis tools applied.

Section 6.2 has discussed that one limitation of this study is the length of innovation conducted. It is suggested that a more appropriate length of introducing the HOTS approach would be one academic year with the retention test being done at the end of the second year. In university L1 classrooms a significant effect of using high cognitive questions on thinking could be shown after one semester's training with the retention

tested at the end of the second semester (Yang et al. 2005). However, in L2 classrooms the language barrier needs to be taken into account. It requires longer for L2 students to adapt to both high cognitive thinking and language use. Although Eisenman and Payne (1997) found that a HOTS programme was more effective after two years of treatment with young adolescent, it could be difficult to conduct a two-year experiment in Taiwanese universities due to the availability of participants. Thus, it was recommended one-year treatment with the retention test done at the end of the second year. The results obtained should yield a better picture of the effectiveness of this approach.

With regard to analysis tools used to determine the length of utterance and grammatical complexity, the main limitation is that both formulae contain two variables, so that difficulties arise when comparing data across pre- and post-test phases and classes. The results gained at different phases were easily affected by a change in either variable. To solve this problem, both variables in the formula, e.g. total number of clauses and AS units, need to be taken into account when evaluating the effect on students' performance or use alternative analysis tools as discussed in the following section.

The analysis tools used in the current study could be considered suitable for a post-test only research design. For a design with pre- and post-test, it is recommended that total number of words be adopted to determine the effect on the length of utterance at different stages (Alcon 1993). The alternative tools for determining syntactical complexity, different verb forms (Yuan and Ellis's 2003) and using categories of phrase, simple sentence and compound sentence (Cole and Williams 1973) are recommended. These methods would allow easier comparison of the results across phases and classes, as there is only one variable involved.

6.6 Significance of this study

This study has represented a new look at L2 speaking problems and educational goals in HE system. It explored the relationship between higher-order thinking and talk in L2 learning. The results have also demonstrated the practicability of using the HOTS approach to develop speaking ability and foster higher-order thinking skills in a Taiwanese university L2 context. This case study of an innovation is the first research of this type to be found outside of westernized countries and former British colonies. It has shown that despite dominant trends that indicate students' opportunities to use English being limited with students' learning style being passive and critical thinking being less practiced, high cognitive thinking abilities can be improved through instruction using the HOTS approach. This contrasts with Atkinson's (1997) argument that it is difficult or impossible to train students with Chinese cultures to think critically in L2 classrooms.

This study has shown the value of thinking tasks which support higher-order questions, providing evidence that not only should the effectiveness of questions be considered, but also the influence they have on students' attitudes. Unlike previous research (Alcon 1993; Ayaduray and Jacobs 1997) which explored only the effect on speaking and thinking performance, this study has investigated both the effect and students' attitudes including student disposition towards speaking and learning and the complex relationship between speaking and attitudes. Thus, this study reveals a more in-depth and comprehensive understanding towards the use of HOTS in L2 learning.

The findings have shown that high cognitive interactive talk provides appropriate speaking and thinking opportunities for effective learning. Such high cognitive interaction allows students to use high cognitive thinking and try out their linguistic hypotheses, thus L2 speaking and thinking can be enhanced simultaneously. Very importantly, students themselves can further create numerous learning opportunities.

The HOTS approach can generate wider-ranging effects than tasks (for example, Foster and Skehan 1996; Yuan and Ellis 2003) which focus only on linguistic development. It brings more benefit for L2 learning.

The significance of the impact on students' speaking performance lies in a good command of spoken English. Previous research (Alcon 1993; Godfrey 2001) investigated only students' linguistic competence; this study has further explored students' speaking proficiency as well. Syntactic complexity does not necessarily have a positive relationship with L2 proficiency. The core issue for learning L2 speaking is to use it for communication. Thus, on top of good linguistic competence, it is important for students to be able to manipulate the target language in a social context. This study has found that the HOTS approach can enable students to better grasp real-life principles of language.

In association, this study has found that the HOTS approach enables students not only to be more cognitively active, but also to use these thinking skills for interaction. Previous studies (Ayaduray and Jacobs 1997; Godfrey 2001) explored only the frequency of HOTS, the present study has further shown that the HOTS approach can facilitate the ability to use high cognition in a social context. It can be argued that merely answering higher-order questions without any exchange of views is seen as passive in terms of using high cognition. It is through operating these thinking skills for communication, a genuine ability of thinking proactively and independently shows.

Importantly the findings have shown that the effect on students' speaking and thinking performance can be maintained where previous research (Alcon 1993; Ayaduray and Jacobs 1997) has only explored the immediate impact. As Mackey and Gass (2005) argue, the real question for studies of second language learning is to address the extent to which a treatment truly resulted in learning. The present study has addressed

this issue and found that the HOTS approach not only has an immediate effect on student performance, but the effect is maintained in the majority of students.

The HOTS approach requires students to think hard and such hard work arguably places students under pressure. Yet, the majority of students still have positive attitudes towards using high cognition in L2 learning because of the benefit this approach carries. As Tan (2007) found that simply conducting higher-order questioning has negative affective impact on students. This study has indicated that incorporating the use of higher-order questioning and the thinking tasks offers a greater opportunity for students to adapt to such thinking approach and the hard work associated.

The HOTS approach associated with social constructivism is not the sole prerogative of more capable learners, but is expected at all levels. It benefited not only high achievers but also low achievers. This implies that this approach may have the valuable effect of reducing the educational disadvantage of low achievers while also being beneficial for high achievers.

This study has shown that learning contexts emerged from the HOTS approach provide solutions to L2 speaking problems observed in Taiwan. It tackles problems of material design, learning context, low motivation, negative dispositions and low speaking proficiency. Most importantly, it carries out the educational goal of Taiwanese Higher Education which is to support students to become active thinkers and use high cognition while learning in L2 classrooms. Overall, this study has provoked thinking and reflection which could lead towards a re-conceptualisation of L2 teaching and learning in Taiwanese HE.

BIBLIOGRAPHY

Adams-Smith, D., 1981. Levels of questioning: teaching creative thinking through ESP. English teaching forum, 19, 15-17

Adey, P., 2004. The professional development of teachers: practice and theory. London: Kluwer Academic.

Alcon, E. G., 1993. High cognitive question in NNS group classroom discussion: do they facilitate comprehension & production of the foreign language? RELC Journal, 24, 73-85

Alexander, R., 2003. Talk in teaching and learning: international perspectives. *In.* New perspectives on spoken English in the classroom. London: QCA.

Ali, A. M. and Daud, N. M., 2003. Using high cognitive questions in ESL classroom to develop critical thinking skills. *In.* A. Pandian, G. Chakravarthy and S. C. Lah, eds. English language teaching and literacy: Research and Reflection. Sefdang: University Putra Malaysia, 165-174.

Allwright, R., 1979. Language learning through communication practice. *In.* C. J. Brumfit and K. Johnson, eds. The communicative approach to language teaching. Oxford: Oxford University, 167-182.

Arnold, D. S., Atwood, R. K. and Rogers, V. M., 1974. Question and response levels and lapse time intervals. Journal of experimental education, 43, 11-15

Atkinson, D., 1997. A critical approach to critical thinking in TESOL. A critical approach to critical thinking in TESOL. TESOL quarterly, 31 (1), 71-94

Ayaduray, J. and Jacobs, G. M. (1997). Can learner strategy instruction succeed? The case of higher order questions and elaborated responses. *System* [online], 25 (4). Available from: http://www.sciencedirect.com/science?_ob=MImg&_imagekey=B6VCH-3SWSGM1-7-2&_cdi=5955&_user=224739&_orig=search&_coverDate=12%2F31%2F1997&_sk=999749995&view=c&wchp=dGLzVzz-zSkWA&md5=46d4a4ac65f2f485cf23733ddc8be417&ie=/sdarticle.pdf [Accessed 8 Jul 2007].

Babbie, E., 2004. The practice of social research. 10th ed. London: Thomson/Wadsworth.

Barden, M. B. (1995). Effective questioning and the ever-elusive higher-order questions. *The American biology teacher* [online], 57 (7). Available from: http://www.jstor.org/stable/pdfplus/4450031.pdf [Accessed 8 Nov 2007].

Bearison, D. J., Magzamen, S. and Filardo, E. K., 1986. Socio-cognitive conflict and cognitive growth in young children. Merrill-Palmer Quarterly, 32 (1), 51-72

Benesch, S., 1999. Thinking critically, thinking dialogically. TESOL Quarterly, 33 (3), 573-580

Bentley, T., 2000. Learning beyond the classroom: education for a changing world. London: Routledge.

Berg, B. L., 2006. Qualitative research methods for the social sciences. 6th ed. New York: Allyn and Bacon.

Billing, D. (2007). Teaching for transfer of core/key skills in higher education: cognitive skills. *Higher education* [online], 53 (4). Available from: http://www.springerlink.com/content/620575741416l651h5/fulltext.pdf [Accessed 3 Aug 2009].

Bloom, B. S., 1956. Taxonomy of educational objectives: the classification of educational goals. New York: Longman.

Blosser, P. E., 1973. Handbook of effective questioning techniques. Worthington, Ohio: Education Associates.

Bowdler, D., Webb, T. and Dyke, S., 1992. Using 'Somerset thinking skills' to promote oracy in the classroom. *In.* N. Jones and N. Jones, eds. Learning to behave: curriculum and whole school management approaches to discipline. London: Kogan, 154-169.

Brock, C. A., 1986. The effect of referential questions on ESL classroom discourse. TESOL quarterly, 20, 47-58

Brown, G. and Yule, G., 1983. Discourse analysis. Cambridge : Cambridge University.

Brown, J. D., 1989. Language program evaluation: a synthesis of existing possibilities. *In.* K. Johnson, ed. The second language curriculum. Cambridge: Cambridge University, 222-241.

Bruner, J. S., 1972. The relevance of education. London: Allen and Unwin.

Bruner, J. S., 1985. Vygotsky: an historical and conceptual perspective. *In.* V. James, ed. Culture, communication, and cognition: Vygotskian perspectives. London: Cambridge University, 31-32.

Bruner, J. S., 1986. Actual minds, possible worlds. Cambridge, MA: Harvard University.

Bryman, A., 2001. Social research methods. Oxford: Oxford University.

Bryman, A., 2004. Social research methods. 2nd ed. Oxford: Oxford University.

Bryman, A. (2007). Barriers to integrating quantitative and qualitative research. *Journal of mixed methods research* [online], 1 (1). Available from: http://mmr.sagepub.com/cgi/reprint/1/1/8 [Accessed 16 Sep 2008].

Burke, L. A. and Williams, J. M., 2008. Developing young thinkers: an intervention aimed to enhance children's thinking skills. Thinking skills and creativity, 3, 104-124

Burns, A. and Joyce, H., 1997. Focus on speaking. North Ryde, N. S. W.: National Centre for English Language Teaching and Research.

Butterworth, M. and O'Connor, M., 2005. Thinking through English. Cambridge: Chris Kington.

Bygate, M., 1999. Quality of language and purpose of task: patterns of learners' language on two oral communication tasks. Language teaching research, 3 (3), 185-214

Cai, Y., 2003. Classroom interaction: an analysis and evaluation of EFL teachers' question in China's setting. Thesis (MA). Lancaster University.

Cam, P., 1995. Thinking together. Sydney: Primary English teaching association.

Carlsen, W. S., 1991. Questioning in classrooms: a sociolinguistic perspective. Review of educational research, 61 (2), 157-178

Carrell, P. L., 1987. Content and formal schemata in ESL reading. TESOL quarterly, 21, 461-481

Casanave, C. P. (2010). Taking risks? A case study of three doctoral students writing qualitative dissertations at an American university in Japan. *Journal of second language writing* [online], 19 (1). Available from: http://www.sciencedirect.com/science?_ob=MImg&_imagekey=B6W5F-4Y4XCW6-1-1&_cdi=6569&_user=224739&_pii=S1060374309000514&_orig=browse&_coverDate=03%2F31%2F2010&_sk=999809998&view=c&wchp=dGLzVtb-zSkzk&md5=7e0e7b7f6eb86f547fac99720140d52a&ie=/sdarticle.pdf [Accessed 2 Mar 2010].

Chaudron, C., 1988. Second language classrooms: research on teaching and learning. Cambridge: Cambridge University.

Chen, Y., 2002. The relationship between anxiety and English proficiency of EFL learners in Taiwan. Thesis (MA). National Kaohsiung First University of Science and Technology.

China Times e-paper (2009). English village [online]. Available from http://tw.news.yahoo.com/article/url/d/a/090421/4/1i69s.html [Accessed 21 April 2009].

Chou, H., 2006. The relationship between critical thinking skills and critical thinking disposition of college students in Taiwan. Thesis (MA). National Cheng-Kung University.

Clément, R., Dörnyei, Z. and Noels, K. A., 1994. Motivation, self-confidence, and group cohesion in the foreign language classroom. Language learning, 44 (3), 417-448

Coffield, F., Moseley, D., Hall, E. and Ecclestone, K. (2004). Should we be using learning styles? what research has to say to practice. *Learning & skills research centre* [online]. Available from: https://crm.lsnlearning.org.uk/user/login.aspx?code=041540&P=041540PD&action=pdfdl&src=xoweb [Accessed 28 Feb 2010].

Cohen, L. and Manion, L., 1994. Research methods in education. 4th ed. London: Routledge.

Cohen, L., Manion, L. and Morrison, K., 2007. Research methods in education. 6th ed. London: Routledge.

Cole, R. A. and Williams, D. M., 1973. Pupil responses to teacher questions: cognitive level, length and syntax. Educational leadership, 31, 142-145

Cook, V., 1986. The basis for an experimental approach to second language learning. *In.* V. Cook, ed. Experimental approaches to second language learning. Oxford: Pergamon.

Costa, A. L. ed., 2001. Developing minds: a resource book for teaching thinking. 3rd ed. Alexandris, VA.: Association for supervision and curriculum development.

Coyle, D., 2002. Towards a reconceptualisation of the MFL curriculum. *In.* A. Swarbrick, ed. Teaching modern foreign languages in secondary schools: a reader. London: Routledge, 156-172.

Crabtree, B. F. and Miller, W. L., 1999. Using codes and code manuals: . A template organising style of interpretation. *In.* B. F. Crabtree and W. L. Miller, eds. Doing qualitative research. London: Sage, 163-177.

Cresswell, J. W., 2003. Research design: qualitative, quantitative, and mixed methods approaches. 2nd ed. Thousand Oaks: Sage.

Currall, S. C. and Towler, A. J., 2002. Research methods in management and organizational research: toward integration of qualitative and quantitative techniques. *In.* A. Tashakkori and C. Teddlie, eds. Handbook of mixed methods in social and behavioral research. London: Sage, 513-526.

Day, R., 2002. Critical thinking in the EFL classroom. *In:* 11th international symposium on English teaching/4th pan Asian conference, Taipei 8-10 November 2002. Taipei: Crane, 72-78.

De Bono, E., 2000. Six thinking hats. London: Penguin.

De Vaus, D., 2001. Research design in social research. London: Sage.

Denzin, N. K., 1978. The research act: a theoretical introduction to sociological methods. New York: Praeger.

Detterman, D. K., 1993. The case for the prosecution: transfer as an epiphenomenon. *In.* D. K. Detterman and R. J. Sternberg, eds. Transfer on Trial: intelligence, cognition, and instruction. Norwood: Ablex, 1-24.

Dillon, J. T., 1978. Using questions to depress student thought. School review, 87, 50-63

Dillon, J. T., 1979. Alternatives to questioning. High school journal, 62, 217-222

Dillon, J. T. (1981). Duration of response to teacher questions and statements. *Contemporary educational psychology* [online], 6 (1). Available from: http://www.sciencedirect.com/science?_ob=MImg&_imagekey=B6WD1-4D6RJHD-8P-1&_cdi=6753&_user=224739&_orig=browse&_coverDate=01%2F31%2F1981&_sk=999939998&view=c&wchp=dGLbVlW-zSkzk&md5=6f91333bd9f929069fe2ad29971ed963&ie=/sdarticle.pdf [Accessed 10 Jul 2007].

Ding, Y. (2007). Text memorization and imitation: the practices of successful Chinese learners of English. *System* [online], 35. Available from: http://www.sciencedirect.com/science?_ob=MImg&_imagekey=B6VCH-4NHM6KN-1-1&_cdi=5955&_user=224739&_orig=search&_coverDate=06%2F30%2F2007&_sk=999649997&view=c&wchp=dGLbVlW-zSkWb&md5=d16ea06c833b60d3ca7b5837e3eeef86&ie=/sdarticle.pdf [Accessed 2 Nov 2009].

Doise, W. and Mugny, G., 1979. Individual and collective conflicts of centrations in cognitive development. European journal of social psychology, 9, 105-108

Donato, R., 1994. Collective Scaffolding in second Language Learning. *In.* J. Lantolf and G. Appel, eds. Vygotskian approaches to second language research. Westport, Conn.: Ablex, 33-56.

Dörnyei, Z., 2003. Questionnaires in second language research: construction, administration, and processing. Mahwah, N.J.: Lawrence Erlbaum.

Duell, O. K., Lynch, D. J., Ellsworth, R. and Moor, C. A., 1992. Wait-time in college classes taken by education majors. Research in higher education, 33 (4), 483-495

Dunkin, M. J. and Biddle, B. J., 1974. The study of teaching. New York: Holt, Rinehart, & Winston.

Eisenman, G. and Payne, B. D., 1997. Effects of the higher order thinking skills program on at-risk young adolescents' self-concept, reading achievement, and thinking skills. Research in middle level education quarterly, 20 (3), 1-25

Ellis, R., 1990. Instructed second language acquisition: learning in the classroom. Oxford: Blackwell.

Ellis, R., 2005. Analysing learner language. Oxford: Oxford University.

Emerson, C. and Holquist, M., 1981. Glossary. *In.* M. Holquist, ed. The dialogic imagination. Austin: University of Texas, 423-434.

Ennis, R. R., 1989. Critical thinking and subject specificity: clarification and needed research. Educational researcher, 18, 4-10

Fairclough, N., 1989. Language and power. London: Longman.

Feuerstein, R., Rand, Y., Hoffman, M. B. and Miller, R., 1979. Instrumental enrichment: an intervention program for cognitive modifiability. Baltimore: University Park.

Field, A., 2000. Discovering statistics using SPSS for windows: advanced techniques for the beginner. London: Sage.

Fine, G., Morrill, C. and Surianarain, S., 2009. Ethnography in organizational settings. *In.* D. Buchanan and A. Bryman, eds. The SAGE handbook of organizational research methods. London: SAGE, 602-619.

Fisher, R., 1998. Teaching thinking: philosophical enquiry in the classroom. London: Cassell.

Foote, C. J., 1998. Student-generated higher order questioning as a study strategy. Journal of educational research, 92 (2), 107-113

Foster, P. and Skehan, P., 1996. The influence of planning and task type on second language performance. Studies in second language acquisition, 18 (3), 299-323

Foster, P., Tonkyn, A. and Wigglesworth, G., 2000. Measuring spoken language: a unit for all reasons. Applied linguistics, 21 (3), 354-375

Fox, H., 1994. Listening to the world: cultural issues in academic writing. Urbana, IL: National council of teachers of English.

Franklin, A. J., 1985. The social context and socialization variables as factors in thinking and learning. *In.* S. F. Chipman, J. W. Segal and R. Glaser, eds. Thinking and learning skills. 2nd ed. Hillsdale, NJ: Lawrence Erlbaum, 81-106.

Fredricks, J., Blumenfeld, P. and Paris, A., 2004. School engagement: potential of the concept, state of the evidence. Review of educational research, 74 (1), 59-109

Furby, L. and Beyth-Marom, R., 1992. Risk taking in adolescence: a decision-making perspective. Developmental review, 12, 1-44

Gall, M. D., Borg, W. R. and Gall, J. P., 1996. Educational research: an introduction. 6th ed. New York: Longman.

Gall, M. D., Ward, B. A., Berliner, D. C., Cahen, L. S., Winne, P. H., Elashoff, J. A. and Stanton, G. C., 1978. Effects of questioning techniques and recitation on student learning. American educational research journal, 15, 175-199

Gall, M. D., 1970. The use of questions in teaching. Review of educational research, 40 (5), 707-721

Gass, S. M. and Varonis, E. M., 1994. Input, interaction, and second language production. Studies in second language acquisition, 16 (3), 283-302

Gass, S. M., 1997. Input, interaction and the second language learner. Mahwah, N.J.: Lawrence Erlbaum.

Gieve, S., 1998. Comments on Dwight Atkinson's 'A critical approach to critical thinking in TESOL': another reader reacts. TESOL Quarterly, 32 (1), 123-129

Gilbert, G. N. ed., 2008. Researching social life. 3rd ed. London: Sage.

Givens, C. G., 1976. A descriptive study of the cognitive level of classroom discourse of college professors and students. Thesis (PhD). Claremont Graduate School.

Glaser, R., 1984. Education and thinking: the role of knowledge. American psychologist, 39, 93-104

Godfrey, K. A. (2001). *Teacher questioning techniques, student responses and critical thinking* [online]. Thesis (MA). Portland State University. Available from: http://www.eric.ed.gov/ERICDocs/data/ericdocs2sql/content_storage_01/0000019b/80/19/8d/27.pdf [Accessed 3 Oct 2007].

Gokhale, A. A., 1995. Collaborative learning enhances critical thinking. Journal of technology education, 7 (1), 22-30

Green, J. C., Caracelli, V. J. and Graham, W. F., 1989. Towards a conceptual framework for mixed-method evaluation designs. Educational evaluation and policy analysis, 11, 255-274

Green, J. L. and Harker, J. O. eds., 1988. Multiple perspective analysis of classroom discourse. Norwood, NJ: Ablex.

Green, J. M., 1993. Learner attitudes toward communicative and non-communicative activities: do enjoyment and effectiveness go together? The modern language journal, 77 (1), 1-10

Halpern, D. F. (1999). Teaching for critical thinking: Helping college students develop the skills and dispositions of a critical thinker. *New directions for teaching and learning* [online], 80. Available from: http://www3.interscience.wiley.com/cgi-bin/fulltext/101523939/PDFSTART [Accessed 22 Aug 2007].

Halx, M. D. and Reybold, L. E. (2006). A pedagogy of force: faculty perspectives of critical thinking capacity in undergraduate students. *The journal of general education* [online], 54 (4). Available from: http://muse.jhu.edu/journals/journal_of_general_education/v054/54.4halx.pdf [Accessed 23 Sep 2007].

Hammersley, M., 1996. The relationship between qualitative and quantitative research: paradigm loyalty versus methodological eclecticism. *In.* J. T. E. Richardson, ed. Handbook of qualitative research methods for psychology and the social sciences. Leicester: BPS Books, 159-174.

Higgins, S. and Baumfield, V., 2001. Thinking through primary teaching. Cambridge: Chris Kington.

Higgins, S., Hall, E., Baumfield, V. and Moseley, D. (2005). *A meta-analysis of the impact of the implementation of thinking skills approaches on pupils* [online]. London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London. Available from: http://eppi.ioe.ac.uk/cms/LinkClick.aspx?fileticket=8%2b7zdZltqAg%3d&tabid=339&mid=1262&language=en-US [Accessed 7 Nov 2007].

Ho, Y., 2002. Comparing the effects of three types of drama activities on the oral practice of technology students. *In:* 11th international symposium on English teaching/ 4th pan Asian conference. Taipei 8-10 November 2002. Taipei: Crane, 336-347.

Holliday, A., 1994. Appropriate methodology and social context. Cambridge: Cambridge University.

Holsti, O. R., 1969. Content analysis for the social sciences and humanities. Reading, MA: Addison-Wesley.

Honea, J. M. Jr., 1982. Wait-time as an instructional variable: an influence on teacher and student. Clearing house, 56 (4), 167-170

Howe, K. R., 1988. Against the quantitative-qualitative incompatability thesis or dogmas die hard. Educational researcher, 17, 10-16

Hsu, L., 2004. The teaching of speaking in the language classroom form a discourse perspective. *In:* 13th international symposium on English teaching/ 6th pan Asian conference. Taipei 11-13 November 2004. Taipei: Crane, 87-94.

Huang, S. Y., 1998. Are freshman English programs for non-English majors at universities in Taiwan meeting the expectation of students and teachers?, National Science council (ROC) project report, 9.2. Taipei, Republic of China: National Science council press, 201-211.

Hughes, J. and Kwok, O., 2007. Influence of student-teacher and parent-teacher relationships on lower achieving readers' engagement and achievement in the primary grades. Journal of educational psychology, 99 (1), 39-51

Hughes, M., 2002. Interviewing. *In.* T. Greenfield ed. Research Methods for Postgraduates. 2nd ed. London: Arnold, 209-217.

Hunt, K., 1965. Grammatical structures written at three grade levels. Champaign, ILL: National council of teachers of English.

Hunter-Grundin, E., 1985. Teaching thinking: an evaluation of Edward de Bono's classroom material. London: Schools Council.

James, M. A. (2006). Teaching for transfer in ELT. *ELT journal* [online], 60 (2). Available from: http://eltj.oxfordjournals.org/cgi/content/abstract/60/2/151 [Accessed 10 Oct 2009].

Johnson, D. M., 1992. Approaches to research in second language learning. New York: Longman.

Johnson, D. W., Johnson, R. T. and Holubec, E. J., 1993. Circles of learning. 4th ed. Edina, MN: Interaction Books.

Johnson, R. B. and Onwuegbuzie, A. J. (2004). Mixed methods research: a research paradigm whose time has come. *Educational researcher* [online], 33 (7). Available from: http://edr.sagepub.com/cgi/reprint/33/7/14 [Accessed 9 Sep 2008].

Johnson, R. T. and Johnson, D. W., 1986. Action research: cooperative learning in the science classroom. Science and children, 24, 31-32

Jones, B. F., Palincsar, A. S., Ogle, D. S. and Carr, E. G. eds., 1987. Strategic teaching and learning: cognitive instruction in the content areas. Alexandria, VA: Association for supervision and curriculum development.

Kagan, S., 1994. Cooperative learning. San Juan Capistrano, CA: Kagan Cooperative Learning.

Kennedy, M., Fisher, M. and Ennis, R. H., 1991. Critical thinking: Literature review and needed research. *In.* L. Idol and B. F. Jones, eds. Educational values and cognitive instruction: implications for reform. Hillsdale, NJ: Lawrence Erlbaum, 11-40.

King, A. (1989). Effects of self-questioning training on college students' comprehension of lectures. *Contemporary educational psychology* [online], 14 (4). Available from: http://www.sciencedirect.com/science?_ob=MImg&_imagekey=B6WD1-4D6RK51-GR-1&_cdi=6753&_user=224739&_orig=browse&_coverDate=10%2F31%2F1989&

sk=999859995&view=c&wchp=dGLbVzb-zSkzV&md5=cfb3276c6a66f47347081ffe887641da&ie=/sdarticle.pdf [Accessed 12 Sep 2007].

King, A., 1990. Enhancing peer interaction and learning in the classroom through reciprocal questioning. American educational research journal, 27 (4), 664-687

Kissock, C. and Iyortsuun, P. T., 1982. A guide to questioning: classroom procedures for teachers. London: Macmillan.

Kowal, M. and Swain, M., 1994. Using collaborative language production tasks to promote students' language awareness. Language awareness, 3, 73-93

Kowal, M. and Swain, M., 1997. From semantic to syntactic processing: how can we promote it in the immersion classroom? *In.* K. Johnson and M. Swain, eds. Immersion education: international perspectives. Cambridge: Cambridge University, 284-309.

Krupa-Kwiatkowski, M., 1998. You shouldn't have brought me here: interaction strategies in the silent period of an inner-direct second language learner. Research on language and social interaction, 31 (2), 133-175

Kuhn, D., 2005. Education for thinking. London: Harvard University.

Lake, M. and Needham, M., 1995. Top ten tactics. Birmingham, England: Questions.

Lantolf, J. P. and Appel, G. eds., 1994. Vygotskian approaches to second language research. Norwood, NJ: Ablex.

Leat, D., 1998. Thinking through geography. Cambridge: Chris Kington.

Legard, R., Keegan, J. and Ward, K., 2003. In-depth interviews. *In.* J. Ritchie and J. Lewis, eds. Qualitative research practice: a guide for social science students and researchers. London: Sage, 138-169.

Legutke, M. and Thomas, H., 1991. Process and experience in the language classroom. London: Longman.

Leith, J., 2000. The chattering classes. *In.* TES curriculum special: modern languages times education supplement, 11-12. Lisbon European council presidency conclusion, March 2000.

Liao, M., 2008. Comparison of English culture apprehension and oral English apprehension in English counseling for English majors and non-English majors. Thesis (MA). National Kaohsiung Normal University.

Likert, R., 1932. A technique for the measurement of attitudes. Archives of psychology, 140, 1-55

Lin, C., 2005. Teaching critical thinking through films in the EFL classroom. *In:* 14th international symposium on English teaching. Taipei 11-13 November 2005. Taipei: Crane, 185-193.

Lin, M., Chen, M. H. and Bedir, H., 2010. Teacher education: Infusing thinking skills into ELT curriculum. *In*: 44th annual international iatefl conference and exhibition, Harrogate 7-11 Apr 2010.

Lin, M. and Mackay, C., 2004. Thinking through modern foreign languages. Cambridge: Chris Kington.

Lipman, M., Sharp, A. M. and Oscanyan, F. S., 1980. Philosophy in the classroom. 2nd ed. Philadelphia: Temple University.

Littlewood, W., 2000. Do Asian students really want to listen and obey? ELT journal, 54 (1), 31-36

Liu, D., 1998. Ethnocentrism in TESOL: teacher education and the neglected needs of international TESOL students. ELT journal, 52 (1), 3-10

Liu, G. (2005). The trend and challenge for teaching EFL at Taiwanese universities. *RELC journal* [online], 36 (2). Available from: http://rel.sagepub.com/cgi/reprint/36/2/211 [Accessed 22 Dec 2008].

Loban, W., 1966. The language of elementary school children. Research report no. 1. Champaign, Ill: National council of teachers of English.

Long, M. H. and Porter, P. A., 1985. Group work, interlanguage talk and second language acquisition. TESOL quarterly, 19, 207-228

Long, M. H. and Sato, C., 1983. Classroom foreigner talk discourse: forms and functions of teachers' questions. *In.* H. Seliger and M. Long, eds. Classroom-oriented research on second language acquisition. Rowley, Mass.: Newbury House, 268-286.

Long, M. H., 1981. Input, interaction and second language acquisition. *In.* H. Winitz, ed. Native language and foreign language acquisition. Annals of the New York academy of sciences, 379.

Long, M. H., 1983. Linguistics and conversational adjustments to non-native speakers. Studies in second language acquisition, 5, 177-193

Long, M. H., 1985. Input and second language acquisition theory. *In.* S. M. Gass and C. G. Madden, eds. Input in second language acquisition. Rowley, MA: Newbury House, 377-393.

Long, M. H., 1996. The role of the linguistic environment in second language acquisition. *In.* W. C. Ritchie and T. K. Bhatia, eds. Handbook of second language acquisition. New York: Academic, 413-468.

Loschky, L. C., 1994. Comprehensible input and second language acquisition: what is the relationship? Studies in second language acquisition, 16 (3), 303-325

MacIntyre, P. D., Clément, R., Dörnyei, Z. and Noels, K. A., 1998. Conceptualizing willingness to communicate in a L2: a situational model of L2 confidence and affiliation. The modern language journal, 82 (4), 545-562

Mackey, A. (1999). Input, interaction and second language development: an empirical study of question formation in ESL. *Studies in second language acquisition* [online], 21 (4). Available from: http://journals.cambridge.org/download.php?file=%2FSLA%2FSLA21_04%2FS0272263199004027a.pdf&code=07740a70fe47a855d990e75cf703ee45 [Accessed 11 Nov 2008].

Mackey, A. and Gass, S. M., 2005. Second language research: methodology and design. Mahwah, NJ: Lawrence Erlbaum.

Madey, D. L., 1982. Some benefits of integrating qualitative and quantitative methods in program evaluation, with illustrations. Educational evaluation and policy analysis, 4, 223-236

Mantero, M. (2002). Bridging the gap: discourse in text-based foreign language classroom. *Foreign language annals* [online], 35 (4). Available from: http://www3.interscience.wiley.com/cgi-bin/fulltext/121507904/PDFSTART [Accessed 21 Oct 2008].

Marini, A. and Genereux, R., 1995. The challenge of teaching for transfer. In. A. McKeough, J. L. Lupart, and A. Marini, eds. Teaching for transfer: fostering generalization in learning. New Jersey: Lawrence Erlbaum, 1-20.

Marzano, R. J., Brandt, R. S., Hughes, C. S., Jones, B. F., Presseisen, B. Z., Rankin, S. C. and Schor, C., 1988. Dimensions of thinking: a framework for curriculum and instruction. Alexandria, VA: Association for supervision and curriculum development.

McGregor, D., 2007. Developing thinking developing learning: a guide to thinking skills in education. Maidenhead, Berkshire, England: Open University.

McGuinness, C., 1999. From thinking skills to thinking classrooms: a review and evaluation of approaches for developing pupils' thinking. London: HMSO.

Mercer, N., 1995. The guided construction of knowledge: talk amongst teachers and learners. Clevedon: Multilingual Matters.

Mercer, N., 2000. Words and mind: how we use language to think together. London: Routledge.

Mercer, N., Wegerif, R. and Dawes, L. (1999). Children's talk and the development of reasoning in the classroom. *British educational research journal* [online], 25 (1). Available from:

http://pdfserve.informaworld.com/686957_731197602_746308972.pdf [Accessed 9 Dec 2009].

Meredith, R. A. (1978). Improved oral test scores through delayed response. *Modern language journal* [online], 62. Available from: http://www.jstor.org/stable/pdfplus/324450.pdf [Accessed 30 May 2007].

Met, M., 1998. Curriculum decision-making in content-based language teaching. *In.* C. Cenoz and F. Genesee, eds. Beyond bilingualism: multilingualism and multilingual education. Clevedon: Multilingual Matters, 35-63.

Miles, M. B. and Huberman, A. M., 1994. Qualitative data analysis: an expanded sourcebook. 2nd ed. Thousand Oaks: Sage.

Milgram, S., 1963. Behavioural study of obedience. Journal of abnormal and social psychology, 67, 371-378

Mills, R., Rice, C., Berliner, D. and Rosseau, E. (1980). The correspondence between teacher questions and student answers in classroom discourse. *Journal of experimental education* [online], 48. Available from: http://pao.chadwyck.co.uk/PDF/1253112442473.pdf [Accessed 22 Jul 2007].

Ministry of Education, Taiwan (2001). *Creativity Education White Paper* [online]. Available from: http://www.creativity.edu.tw/modules/wfsection/article.php?articleid=1 [Accessed 4 Jun 2007].

Ministry of Education, Taiwan (2007a). *Project on the enhancement of creative education.* [online]. Available from: http://hss.edu.tw/doc_detail.php?doc_id=1110&plan_title=&class_plan=170 [Accessed 19 July 2008].

Ministry of Education, Taiwan (2007b). *Higher education, technological & vocational education newsletter* [online]. Available from: http://www.news.high.edu.tw/news003/2007030804.asp?c=0400 [Accessed 29 May 2007].

Ministry of Education, Taiwan (2008). *The project on excellence in instruction* [online]. Available from: http://www.edu.tw/files/list/B0039/附件-2 第二期獎勵大學教學卓越計畫 980312-.pdf [Accessed 22 Sep 2009].

Miri, B., David, B. and Uri, Z. (2007). Purposely teaching for the promotion of higher-order thinking skills: a case of critical thinking. *Research in science education* [online], 37 (4). Available from: http://www.springerlink.com/content/2265621540210432/fulltext.pdf [Accessed 15 Aug 2008].

Morgan, N. and Saxton, J., 1991. Teaching, questioning, and learning. London: Routledge.

Morgan, N. and Saxton, J., 1994. Asking better questions: models, techniquest and classroom activities for engaging students in learning. Markham, Ont.: Pembroke.

Moseley, D., Baumfield, V., Elliott, J., Gregson, M., Higgins, S., Miller, J. and Newton, D., 2005. Frameworks for thinking: a handbook for teaching and learning. Cambridge: Cambridge University.

Moseley, D., Baumfield, V., Higgins, S., Lin, M., Miller, J., Newton, D. and Robson, S., 2004. Thinking skill frameworks for post-16 learners: an evaluation. Learning and Skills Research Center, Newcastle University.

Moslehpour, M. and Chou, C., 2004. Foreign language classroom anxiety and problem solving appraisal of EFL students taught by native and non-native English speaking teachers in Taiwan. *In:* 13th international symposium on English teaching/ 6th pan Asian conference. Taipei 11-13 November 2004. Taipei: Crane, 163-179.

Mugny, G. and Doise, W., 1978. Socio-cognitive conflict and the structure of individual and collective performances. European journal of social psychology, 8, 181-192

Muijs, D., 2004. Doing quantitative research in education with SPSS. London: Sage.

Nagappan, R., 2001. The teaching of higher-order thinking skills in Malaysia. Journal of southeast Asian education, 2 (1), 42-65

Nakatani, Y. (2005). The effects of awareness-raising training on oral communication strategy use. *The modern language journal* [online], 89 (1). Available from: http://www3.interscience.wiley.com/cgi-bin/fulltext/118687936/PDFSTART [Accessed 14 Mar 2008].

National Yunlin University of Science & Technology (2007). *Education aims for universities of science & technology* [online]. Available from: http://bcc.yuntech.edu.tw/index.asp [Accessed 7 Feb 2007].

Newman, I. and Benz, C. R., 1998. Qualitative-quantitative research methodology: exploring the interactive continuum. Carbondale, IL: Southern Illinois University.

Nichols, A. and Kinninment, D., eds, 2001. More thinking through geography. Cambridge: Chris Kington.

Nisbett, R. and Wilson, T., 1977. Telling more than we can know: verbal reports on mental processes. Psychological review, 84, 231-259

Nora, N., 1997. ESL learner difficulties in the Malaysian literature classroom. Thesis (PhD). University of Strathclyde.

Nunan, D. (1987). Communicative language teaching: making it work. *ELT journal* [online], 41 (2). Available from: http://eltj.oxfordjournals.org/cgi/reprint/41/2/136 [Accessed 13 Jun 2007].

Nunan, D., 1990. The questions teachers ask. JALT journal, 12 (2), 187-202

Nunan, D. (2003). The impact of English as a global language on educational policies and practices in the Asia-Pacific region. *TESOL quarterly* [online], 37(4). Available from: http://docserver.ingentaconnect.com/deliver/connect/tesol/00398322/v37n4/s1.pdf?expires=1253116556&id=52067004&titleid=6562&accname=Newcastle+University&checksum=4B37C0AD3FA836839CB58327FBFF7E7E [Accessed 8 Aug 2008].

O'Donnell, A. M. and Dansereau, D. F., 1992. Scripted cooperation in student dyads: a method for analyzing and enhancing academic learning and performance. *In*. R. Hertz-Lazarowitz and N. Miller, eds. Interaction in cooperative groups: the theoretical anatomy of group learning. New York: Cambridge University, 120-144.

Ohta, A. A., 1995. Applying sociocultural theory to an analysis of learner discourse: Learner-learner collaborative interaction in the zone of proximal development. Issues in applied linguistics, 6, 93-122

Onwuegbuzie, A. J. (2000). *Positivists, post-positivists, post-structuralists, and post-modernists: why can't we all get along? Towards a framework for unifying research paradigms* [online]. Available from: http://www.eric.ed.gov/ERICDocs/data/ericdocs2sql/content_storage_01/0000019b/80/16/f7/cd.pdf [Accessed 11 Oct 2008].

Onwuegbuzie, A. J. and Leech, N. L. (2005). On becoming a pragmatic researcher: the importance of combining quantitative and qualitative research methodologies. *International journal of social research methodology* [online], 8 (5). Available from: http://web.ebscohost.com/ehost/pdf?vid=2&hid=103&sid=6b3d3d3f-b198-4fa1-995c-309f6a8e1d00%40sessionmgr11 [Accessed 18 Aug 2008].

Oppenheim, A. N., 1992. Questionnaire design, interviewing and attitude measurement. London: Pinter.

Ortega, L., 2003. Syntactic complexity measures and their relationship to L2 proficiency: a research synthesis of college-level L2 writing. Applied linguistics, 24 (4), 492-518

Oxford, R. L., 1990. Language learning strategies: what every teacher should know. Boston: Heinle & Heinle.

Palincsar, A. S. and Brown, A. L., 1984. Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities. Cognition and instruction, 2, 117-175

Paran, A., 2003. Helping learners to become critical: how coursebooks can help. *In.* W. A. Renandya, ed. Methodology and materials design in language teaching: current perceptions and practices and their implications. Singapore: SEAMEO regional language center, 109-123.

Paul, R. W., 1993. Critical thinking: what every person needs to survive in a rapidly changing world. Santa Rosa, CA: Foundation for critical thinking.

Peers, I. S., 1996. Statistical analysis for education and psychology researchers. London: The Falmer.

Perkins, D. N. and Salomon, G., 1988. Teaching for transfer. Educational leadership, 46, 22-32

Pica, T. and Doughty, C., 1985. The role of group work in classroom second language acquisition. Studies in second language acquisition, 7, 233-248

Pica, T., Halliday, L., Lewis, N. and Morgenthaler, L., 1989. Comprehensible outputs as an outcome of linguistic demands on the learner. Studies in second language acquisition, 11 (1), 63-90

Pogrow, S. (1996). HOTS: helping low achievers in grades 4-8. *Principal* [online]. Available from: http://www.hots.org/article_helping.html [Accessed 27 Oct 2009].

Powell, S. and Makin, M., 1994. Enabling pupils with learning difficulties to reflect on their own thinking. *British educational research journal* [online], 20 (5). Available from: http://pdfserve.informaworld.com/59951_751320496_746313815.pdf [Accessed 11 Dec 2009].

Quicke, J. and Winter, C., 1994. Teaching the language of learning: towards a metacognitive approach to pupil empowerment. *British educational research journal* [online], 20 (4). Available from: http://www.informaworld.com/smpp/content~content=a746313621~db=all [Accessed 11 Dec 2009].

Raimes, A. and Zamel, V., 1997. Response to Ramanathan and Kaplan. Journal of second language writing, 6 (1), 79-81

Ramanathan, V. and Kaplan, R., 1996a. Some problematic 'channels' in the teaching of critical thinking in current L1 composition textbooks: implications for L2 student-writers. Issues in Applied linguistics, 7 (2), 225-249

Ramanathan, V. and Kaplan, R., 1996b. Audience and voice in current L1 composition tests: some implications for ESL student writers. Journal of second language writing, 5 (1), 21-34

Redfield, D. L. and Rousseau, E. W., 1981. A meta-analysis of experimental research on teacher questioning behaviour. Review of educational research, 51 (2), 237-246

Reece, G. J. (2007). Critical thinking and cognitive transfer: implications for the development of online information literacy tutorials. *Research strategies* [online], 20. Available from: http://www.sciencedirect.com/science?_ob=MImg&_imagekey=B6W60-4N6Y5J4-1-1&_cdi=6584&_user=224739&_orig=search&_coverDate=12%2F31%2F2005&_sk=999799995&view=c&wchp=dGLzVzz-zSkzV&md5=111f7898e07b6dd1005cab6a687471af&ie=/sdarticle.pdf [Accessed 18 Jul 2008].

Richards, J. C. and Rogers, T., 2001. Approaches and methods in language teaching. 2nd ed. Cambridge: Cambridge University.

Richards, K., 2003. Qualitative inquiry in TESOL. New York: Palgrave Macmillan.

Ritchhart, R. and Perkins, D. N., 2005. Learning to think: the challenges of teaching thinking. *In*. K. J. Holyoak and R. G. Morrison, eds. The Cambridge handbook of thinking and reasoning. Cambridge: Cambridge University, 775-802.

Robson, C., 2002. Real world research: a resource for social scientists and practitioner-researchers. 2nd ed. Oxford: Blackwell.

Rosenshine, B., 1971. Teaching behaviours and student achievement. London: National foundation for educational research.

Rosenshien, B., 1976. Recent research on teaching behaviours and student achievement. Journal of teacher education, 27, 61-64

Rowe, M., 1974. Pausing phenomena: influence on the quality of instruction. Journal of psycholinguistic research, 3, 203-224

Rowe, M., 1980. Pausing principles and their effects on reasoning in science. New directions for community colleges, 8 (3), 27-34

Rowe, M., 1986. Wait-time: slowing down may be a way of speeding up. Journal of teacher education, 37, 43-50

Ryan, F. L., 1973. Differentiated effects of levels of questioning on students achievement. Journal of experimental education, 41, 63-67

Salomon, G. and Perkins, D. N., 1989. Rocky roads to transfer: rethinking mechanisms of a neglected phenomenon. Educational psychologist, 24 (2), 113-142

Samson, G. E., Strykowski, B., Weinstein, T. and Walberg, H. J., 1987. The effects of teacher questioning levels on student achievement: a quantitative synthesis. Journal of educational research, 80 (5), 290-295

Sanders, N., 1966. Classroom questions: what kinds? New York: Harper & Row.

Santiesteban, A. J., 1976. Teacher questioning performance and student affective outcomes. Journal of research in science teaching, 13, 553-557

Sato, C. J., 1982. Ethnic styles in classroom discourse. *In.* M. Hines and W. Rutherford, eds. On TESOL 81. Washington, DC: TESOL, 11-24.

Sato, C. J., 1988. Origins of complex syntax in interlanguage development. Studies in second language acquisition, 10 (3), 371-395

Saville-Troike, M., 1996. The ethnography of communication. *In.* S. L. McKay and N. H. Hornberger, eds. Sociolinguistics and language teaching. Cambridge: Cambridge University, 351-382.

Scardamalia, M. and Bereiter, C., 1985. Fostering the development of self-regulation in children's knowledge processing. *In.* S. F. Chipman, J. W. Segal, and R. Glaser, eds. Thinking and learning skills: Vol. 2. Research and open questions. Hillsdale, NJ: Erlbaum, 563-577.

Schiffrin, D., 1994. Approaches to discourse. London: Blackwell.

Scollon, R., 1991. Eight legs and one elbow: stance and structure in Chinese English compositions. Paper presented at the International reading Association second North American Conference on adult and adolescent literacy. Banff, Canada March, 1991.

Sechrest, L. and Sidani, S. (1995). Quantitative and qualitative methods: is there an alternative? *Evaluation and program planning* [online], 18 (1). Available from: http://www.sciencedirect.com/science?_ob=MImg&_imagekey=B6V7V-3Y6PG0V-19-2&_cdi=5852&_user=224739&_orig=search&_coverDate=03%2F31%2F1995&_sk=999819998&view=c&wchp=dGLbVlb-zSkzk&md5=31577aa40a3a699dff40bb0b7a7cafad&ie=/sdarticle.pdf [Accessed 10 Sep 2008].

Seliger, H. W., 1977. Does practice make perfect?: a study of interaction patterns and L2 competence. Language learning, 27, 263-278

Shima, C. (2008). Fluidity of peer interaction in a Japanese language classroom. *Electronic journal of foreign language teaching* [online], 5 (1). Available from: http://e-flt.nus.edu.sg/v5sp12008/shima.pdf [Accessed 24 Jan 2009].

Shou, Z., 1995. New era of English teaching. English teaching & learning, 19, 48-56

Shrum, J., 1985. Wait-time and the use of target or native languages. Foreign language annals, 18 (4), 305-310

Sieber, S. D., 1973. The integration of fieldwork and survey methods. The American journal of sociology, 78 (6), 1335-1359

Silverman, D., 1993. Interpreting qualitative data. London: Sage.

Sinclair, J. and Coulthard, M., 1975. Towards and analysis of discourse: the English used by teachers and pupils. London: Oxford.

Sinclair, J. and Coulthard, M., 1992. Towards and analysis of discourse: the English used by teachers and pupils. *In*. M. Coulthard, ed. Advances in spoken discourse analysis. London: Routledge, 1-34.

Sizmur, S. and Osborne, J., 1997. Learning processes and collaborative concept mapping. International journal of science education, 19 (10), 1117-1135

Skehan, P. and Foster, P., 1997. Task type and task processing conditions as influences on foreign language performance. Language teaching research, 1, 185-211

Smith, C. T., 1978. Evaluating answers to comprehension questions. The reading teacher, 31, 896-900

Smith, H. and Higgins, S. (2006). Opening classroom interaction: the importance of feedback. *Cambridge journal of education* [online], 36 (4). Available from: http://pdfserve.informaworld.com/393615_731197602_769627374.pdf [Accessed 17 Nov 2007].

Smith, J. and Paterson, F., 1998. Positively bilingual: classrooms strategies to promote the achievement of bilingual learners. Nottingham: Nottingham Education Authority.

Sousa, D. A., 2001. How the brain learns: a classroom teacher's guide. 2nd ed. Thousand Oaks: Crowin.

Spack, R., 1997. The rhetorical construction of multilingual students. TESOL Quarterly, 31 (4), 765-774

Spear, L. C. and Sternberg, R. J., 1987. Teaching styles: staff development for teaching thinking. Journal of staff development, 8 (3), 35-39

Stahl, R. (1994). *Using 'think-time' and 'wait-time' skillfully in the classroom* [online]. Available from:

http://eric.ed.gov/ERICDocs/data/ericdocs2sql/content_storage_01/0000019b/80/15/a4/86.pdf [Accessed 6 Nov 2007].

Stapleton, P. (2001). Assessing critical thinking in the writing of Japanese university students. *Written communication* [online], 18 (4). Available from: http://wcx.sagepub.com/cgi/reprint/18/4/506 [Accessed 18 Oct 2008].

Stapleton, P. (2002). Critical thinking in Japanese L2 writing: rethinking tired constructs. *ELT journal* [online], 56 (3). Available from: http://eltj.oxfordjournals.org/cgi/content/abstract/56/3/250 [Accessed 29 Apr 2008].

Stevens, R., 1912. The question as a measure of efficiency in instruction: a critical study of classroom practice. Teachers college contributions to education, 48

Sutherland, J. (2006). Promoting group talk and higher-order thinking in pupils by 'coaching' secondary English trainee teachers. *Literacy* [online], 40 (2). Available from: http://www3.interscience.wiley.com/cgi-bin/fulltext/118732203/PDFSTART [Accessed 14 Oct 2009].

Swain, M. and Lapkin, S. (1998). Interaction and second language learning: Two adolescent French immersion students working together. *The Modern language journal* [online], 82. Available from: http://www.jstor.org/stable/pdfplus/329959.pdf [Accessed 9 Oct 2007].

Swain, M. and Lapkin, S. (2002). Talking it through: two French immersion learners' response to reformulation. *International journal of educational research* [online], 37 (3- 4). Available from: http://www.sciencedirect.com/science?_ob=MImg&_imagekey=B6VDF-48XSK29-3-4&_cdi=5981&_user=224739&_pii=S0883035503000065&_orig=search&_coverDate=12%2F31%2F2002&_sk=999629996&view=c&wchp=dGLbVtb-zSkWA&md5=941a7843e0ed986eec4c973bab2f78ef&ie=/sdarticle.pdf [Accessed 16 Mar 2010].

Swain, M., 1985. Communicative competence; some roles of comprehensible input and comprehensible output in its development. *In.* S. M. Gass and C. G. Madden, eds. Input in second language acquisition. Rowley, Mass.: Newbury House, 235-253.

Swain, M., 1995. Three functions of output in second language learning. *In.* G. Cook and B. Seidlhofer, eds. Principle and practice in applied linguistics: studies in honor of H. G. Widdowson. Oxford: Oxford University, 125-144.

Swain, M., 1998. Focus on form through conscious reflection. *In.* C. Doughty and J. Williams, eds. Focus on form in classroom second language acquisition. Cambridge: Cambridge University, 64-81.

Swain, M., 2000. The output hypothesis and beyond: mediating acquisition through collaborative dialogue. *In.* J. Lantolf, ed. Sociocultural theory and second language learning. Oxford: Oxford University, 97-114.

Swain, M., 2005. The output hypothesis: theory and research. *In.* E. Hinkel, ed. Handbook of research in second language teaching and learning. Mahwah, NJ: Erlbaum, 471-483.

Swain, M., 2006. Languaging, agency, and collaboration in advanced second language proficiency. *In.* H. Byrnes, ed. Advanced language learning: the contributions of Halliday and Vygotsky. London: Continuum, 95-108.

SZE, P., 1995. Teaching conversation in the second language classroom: problems and prospects. Educational journal, 23, 229-250

Talebinezahd, M. R., 2003. Effective questions. English teaching forum, 41 (4), 46-47

Tan, Z. (2007). Questioning in Chinese University EL classrooms: what lies beyond it? *RELC journal* [online], 38 (1). Available from: http://rel.sagepub.com/cgi/reprint/38/1/87 [Accessed 25 Nov 2007].

Tanka, J. and Most, P., 2007. Interactions 1 listening/speaking. New York: McGraw-Hill.

Tashakkori, A. and Teddlie, C., 1998. Mixed methodology: combining qualitative and quantitative approaches. Thousand, CA: Sage.

Thinking Skills Review Group, University of London (2004). *Thinking skills approaches to effective teaching and learning: what is the evidence for impact on learners?* [online]. Available from: http://eppi.ioe.ac.uk/cms/Default.aspx?tabid=336 [Accessed 13 Mar 2007].

Thomas, A. (2000). *A mix of chance and challenge* [online]. Available from: http://www.tes.co.uk/article.aspx?storycode=339808 [Accessed 25 Feb 2010].

Thornbury, S., 1999. How to teach grammar. Harlow: Longman.

Thornbury, S., 2000. A dogma for EFL. IATEFL issues, 153, 24-28

Tobin, K., 1987. The role of wait-time in higher cognitive level learning. Review of educational research, 57, 69-95

Totten, S., Sills, T., Digby, A. and Russ, P., 1991. Cooperative learning: a guide to research. New York: Garland.

Tsai, Y. (2007). *What are the implications of the spread of English language education in China and Taiwan?* [online]. Available from: http://www.ced.ncnu.edu.tw/amy/essay5.htm [Accessed 15 Jun 2010].

Tsui, A., 1996. Reticence and anxiety in second language learning. *In.* K. Bailey and D. Nunan, eds. Voices from the language classroom: qualitative research in second language education. Cambridge: Cambridge University, 145-167.

Ustunluoglu, E., 2004. Language teaching through critical thinking and self-awareness. English teaching forum, 42, 2-8

Van Boxtel, C., Van der Linden, J. and Kanselaar, G., 2000. Deep processing in a collaborative learning environment. *In.* H. Cowie and G. Van der Aalsvoort, eds. Social interaction in learning and instruction: the meaning of discourse for the construction of knowledge. Amsterdam: Pergamon, 161-178.

Van Duzer, C. and Florez, M. C. (1999). Critical literacy for adult literacy in language learners. *Washington, DC: National center for ESL literacy education* [online]. Available from: http://www.eric.ed.gov/ERICDocs/data/ericdocs2sql/content_storage_01/0000019b/80/16/38/62.pdf [Accessed 2 Aug 2007].

Vygotsky, L. S., 1978. Mind in society: the development of higher psychological processes. Cambridge, MA: Harvard University.

Vygotsky, L. S., 1986. Thought and language. Cambridge, Massachusetts: MIT.

Walsh, S. (2002). Construction or obstruction: teacher talk and learner involvement in the EFL classroom. *Language Teaching Research* [online], 6 (1). Available from: http://www.swetswise.com/FullTextProxy/swproxy?url=http%3A%2F%2Fltr.sagepub.com%2Fcgi%2Freprint%2F6%2F1%2F3.pdf&ts=1253119612630&cs=3168290164&userName=0126756.ipdirect&emCondId=126756&articleID=29272533&yevoID=1784565&titleID=122645&referer=2&remoteAddr=128.240.229.3&hostType=PRO&swsSessionId=+f2v7O5kdN12dFcbYDPqdA**.pasc2 [Accessed 16 Aug 2007].

Watanabe, Y. and Swain, M. (2007). Effects of proficiency differences and patterns of pair interaction on second language learning: collaborative dialogue between adult ESL learners. *Language teaching research* [online], 11 (2). Available from: http://www.swetswise.com/eAccess/viewAbstract.do?articleID=31653068&titleID=122645 [Accessed 02 Jan 2010].

Waters, A. (2006). Thinking and language learning. *ELT Journal* [online], 60 (4). Available from: http://eltj.oxfordjournals.org/cgi/reprint/60/4/319 [Accessed 8 Nov 2007].

Watson, J. (2000). Social constructivism in the classroom. *Support for learning* [online], 15 (3). Available from: http://www3.interscience.wiley.com/cgi-bin/fulltext/119009952/PDFSTART [Accessed 12 Dec 2008].

Watson, J. (2001). Social constructivism. *Support for learning* [online], 16 (3). Available from: http://www3.interscience.wiley.com/cgi-bin/fulltext/118988473/PDFSTART [Accessed 12 Dec 2008].

Webb, N. (1989). Peer interaction and learning in small groups. *International journal of educational research* [online], 13. Available from: http://www.sciencedirect.com/science?_ob=MImg&_imagekey=B6VDF-4691P2G-7-1&_cdi=5981&_user=224739&_orig=search&_coverDate=12%2F31%2F1989&_sk=999869998&view=c&wchp=dGLbVtz-zSkWz&md5=b366d38c2c62d90dad56356208833832&ie=/sdarticle.pdf [Accessed 11 Nov 2007].

Webb, N. M. and Palincsar, A. S., 1996. Group processes in the classroom. *In.* D. D. B. R. Caffee, ed. Handbook of Educational Psychology. NY: Macmillan, 841-873.

Wegmann, B. and Knezevic, M., 2007. Mosaic 1 reading. New York: Mc Graw-Hill.

Wells, G., 1999. Dialogic inquiry towards a sociocultural practice and theory of education. Cambridge: Cambridge University.

Wenden, A., 1991. Learner strategies for learner autonomy. London: Prentice-Hall International.

Wenden, A., 1997. Designing learner training: the curricular questions. *In.* G. M. Jacobs, ed. Language classrooms of tomorrow: Issues and responses. Regional language center, Singapore: SEAMEO, 238-262.

White, B. Y. and Frederiksen, J. R. (1998). Inquiry, modeling and metacognition: making science accessible to all students. *Cognition and instruction* [online], 16 (1). Available from: http://pdfserve.informaworld.com/987850_751320496_783763026.pdf [Accessed 11 Oct 2009].

White, J. and Lightbown, P. M., 1984. Asking and answering in ESL classes. The Canadian modern language review, 40 (2), 228-244

Wilen, W., 1987. Questioning skills for teachers. Washington, D. C.: National Education Association.

Williams, M., 1991. A framework for teaching English to young learners. *In.* C. Brumfit, J. Moon and R. Tongue, eds. Teaching English to children. London: Collins, 203-212.

Williams, M., 1998. Teaching thinking through a foreign language. *In.* R. Burden and M. Williams, eds. Thinking through the curriculum. London: Routledge, 84-95.

Willson, I. A., 1973. Changes in mean levels of thinking in grades 1-8 through use of an interaction system based on Bloom's taxonomy. Journal of educational research, 66, 13-50

Winne, P. H., 1979. Experiments relating teachers' use of higher cognitive questions to student achievement. Review of educational research, 49 (1), 13-50

Wood, D., Burner, J. and Ross, G., 1976. The role of tutoring in problem solving. Journal of child psychology and psychiatry, 17 (2), 89-100

Wu, K., 1993. Classroom interaction and teacher questions revisited. RELC Journal: a journal of language teaching and research in southeast Asia, 24 (2), 49-68

Xu, J. (2002). *The design and development of freshman English conversation internet-based material* [online]. Thesis (MA). National Taiwan Normal University. Available from: https://etds.ncl.edu.tw/theabs/service/ft_result.jsp?FT=Y&id=090TKU00620001 [Accessed 19 May 2007].

Yamada, S., 1913. A study of questioning. The pedagogical seminary, 20, 129-186

Yang, Y. C., Newby, T. J. and Bill, R. L., 2005. Using Socratic questioning to promote critical thinking skills through asynchronous discussion forums in distance learning environments. The American journal of distance education, 19 (3), 163-181

Yang, Y., 1992. Issues of freshman English teaching. English teaching & learning, 16, 29-37

Yeldham, M., 2001. English conversation in Taiwan (Intermediate level). Taipei: Crane.

Yin, R. K., 2003. Case study research: design and methods. 3rd ed. Thousand Oaks, CA.: Sage.

Yuan, F. and Ellis, R., 2003. The effects of pre-task planning and on-line planning on fluency, complexity and accuracy in L2 monologic oral production. Applied linguistics, 24, 1-27

Yule, G. and Macdonald, D., 1990. Resolving referential conflict in L2 interaction: the effect of proficiency and interactive role. Language learning, 40 (4), 539-556

Zamel, V., 1997. Toward a model of transculturation. TESOL Quarterly, 31 (2), 341-352

Zhong, J., 2001. English instruction material and methodology research for technology colleges. Research project report in national science council.

Zhou, Y. P., 1991. The effect of explicit instruction on the acquisition of English grammatical structures by Chinese learners. *In.* C. James and P. Garrett, eds. Language awareness in the classroom. London: Longman, 254-277.

Zohar, A. and Dori, Y. J. (2003). Higher order thinking skills and low achieving students: are they mutually exclusive? *The journal of the learning sciences* [online], 12 (2). Available from: http://pdfserve.informaworld.com/919549_751320496_785041879.pdf [Accessed 18 Aug 2008].

Zohar, A. and Nemet, F. (2002). Fostering students' knowledge and argumentation skills through dilemmas in human genetics. *Journal of research in science teaching* [online], 39 (1). Available from: http://www3.interscience.wiley.com/cgi-bin/fulltext/88513705/PDFSTART [Accessed 22 Nov 2009].

Zohar, A. (1996). Transfer and retention of reasoning strategies taught in biological contexts. *Research in science and technological education* [online], 14 (2). Available from: http://pdfserve.informaworld.com/791911_751320496_746598199.pdf [Accessed 16 Oct 2009].

Appendix 1 - Thinking Tasks Used in the Innovation

Week 1: 5Ws - Most interesting countries to visit
Thinking skills: analysis, evaluation


France


Italy


Switzerland


Austria


Spain


Germany


Brazil


Greece


America

Task: Choose 3 countries you would most like to visit in your group, and state the reasons.

Countries	Reasons

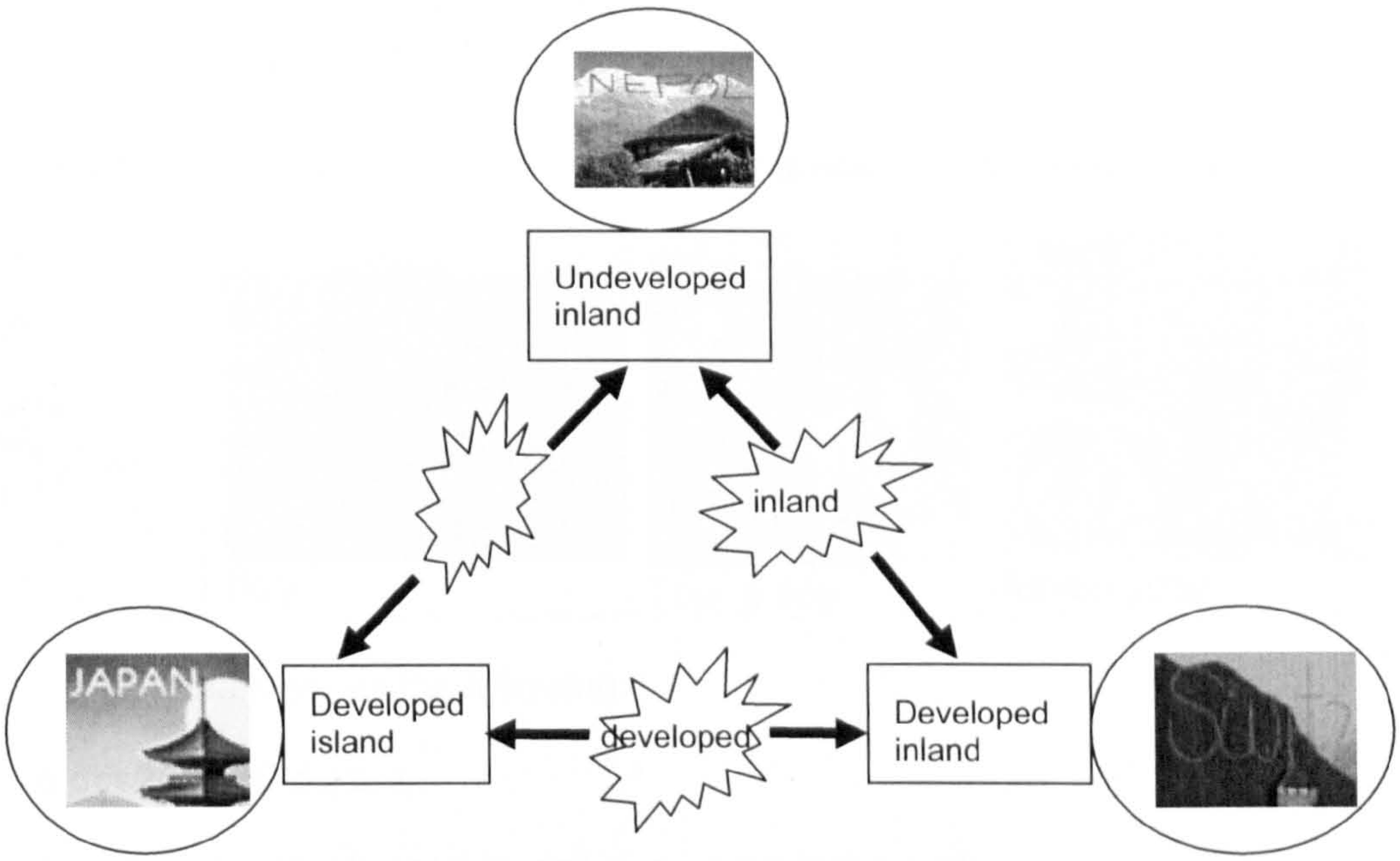
Sentence patterns:

We would most like to visit x because....

The second country we'd like to visit is y because...

The third country we are interested in is z because...

Week 2: Odd One Out - Countries
Thinking skills: analysis (comparing and contrasting)



For example:

1. Nepal is out. Both Japan and Switzerland are developed countries, but Nepal is not.
2. Japan is out. Japan is an island country while Nepal and Switzerland are inland countries.

Week 3: Make Up A Story

Task: Use the following four pictures to Make Up A Story. It is not necessarily to follow the sequence of these pictures. You may reorder the sequence of these pictures.

Thinking skills: synthesis

1. **NT$200,000** 2. **Italy** 3. **Tour guide** 4. **motorcyclist**

			
NT$200,000	Italy	Tour guide	Motorcyclist

1. You may start the story as the following:

   I've got 200,000 dollars…

   Italy is the country I longed to visit and I did it last month …

   I used to work as a tour guide…

   I saw a motorcyclist….

2. Make your own starting sentence.

Week 4: Guess What I Say – Chinese idioms/ proverbs
Thinking skills used by the speaker: comprehension, application
Thinking skills used by the ones who listen and guess: analysis

狐假虎威	藝高人膽大
點石成金	不經一事, 不長一智
不恥下問	秀才遇到兵, 有理說不清
病入膏肓	一回生, 二回熟
騎虎難下	騎驢找馬
食言而肥	真金不怕火煉
無巧不成書	行千里路, 勝讀萬卷書
狐假虎威	藝高人膽大
點石成金	不經一事, 不長一智
不恥下問	秀才遇到兵, 有理說不清
病入膏肓	一回生, 二回熟
騎虎難下	騎驢找馬
食言而肥	真金不怕火煉
無巧不成書	行千里路, 勝讀萬卷書

Take 狐假虎威 (the fox borrows the tiger's fierceness) as an example for language use.

There are four Chinese characters in this idiom. The first and third words are two kinds of animal which are often found in the forest. The first animal often lies to others while the third refers to the king of the forest. The first animal pretends he is the third animal and ....

Week 5: 5Ws – Dog's home

Task:

1. Decide who is suitable to adapt the dog.
2. One dog will be left over. This dog will be destroyed if nobody adopts it. Find a solution to solve the problem.

Thinking skills: synthesis, evaluation, analysis

Visitors to the dogs' home

The Simpsons
Mr. John Simpsons and Mrs. Jane Simpsons, Tony, aged 10 and Lisa, aged 8. They have a house in a quiet street. Everybody goes out to work or school, but they are all home at weekends. The house has a large garden and there is a park 5 minutes away.

Mrs. Robinson
Mrs. Judy Robinson, who is aged 75, lives alone. Her son visits with her two grandchildren, Paul aged 3 and Tina aged 2. She would love a dog for company. She lives in a small home with a very tiny back garden.

The Green Family
Mr. Tim Green and Mrs. Angel Green and Clare, aged 7. Clare has her own pony and two cats. She now wants a dog. Angel is not very keen on dogs. They have a big house in the countryside, which has large garden and is surrounded by fields. They have a Land Rover and enjoy being out-of-doors.

Miss Lee
Miss Pam Lee, aged 31, has a ground floor flat. She works at home on most days. Her hobby is walking and she would like a dog to take with her. She sometimes has her nephew Roy, aged 8, to stay. She does not have a garden.

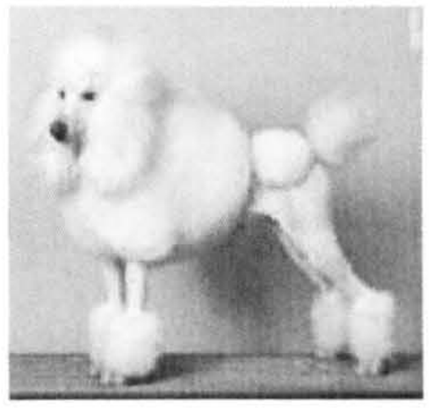
Sentence pattern:

I think Fify can stay with x because...

I don't think Fify can stay with y because...

It's better for Fify to stay with z because...

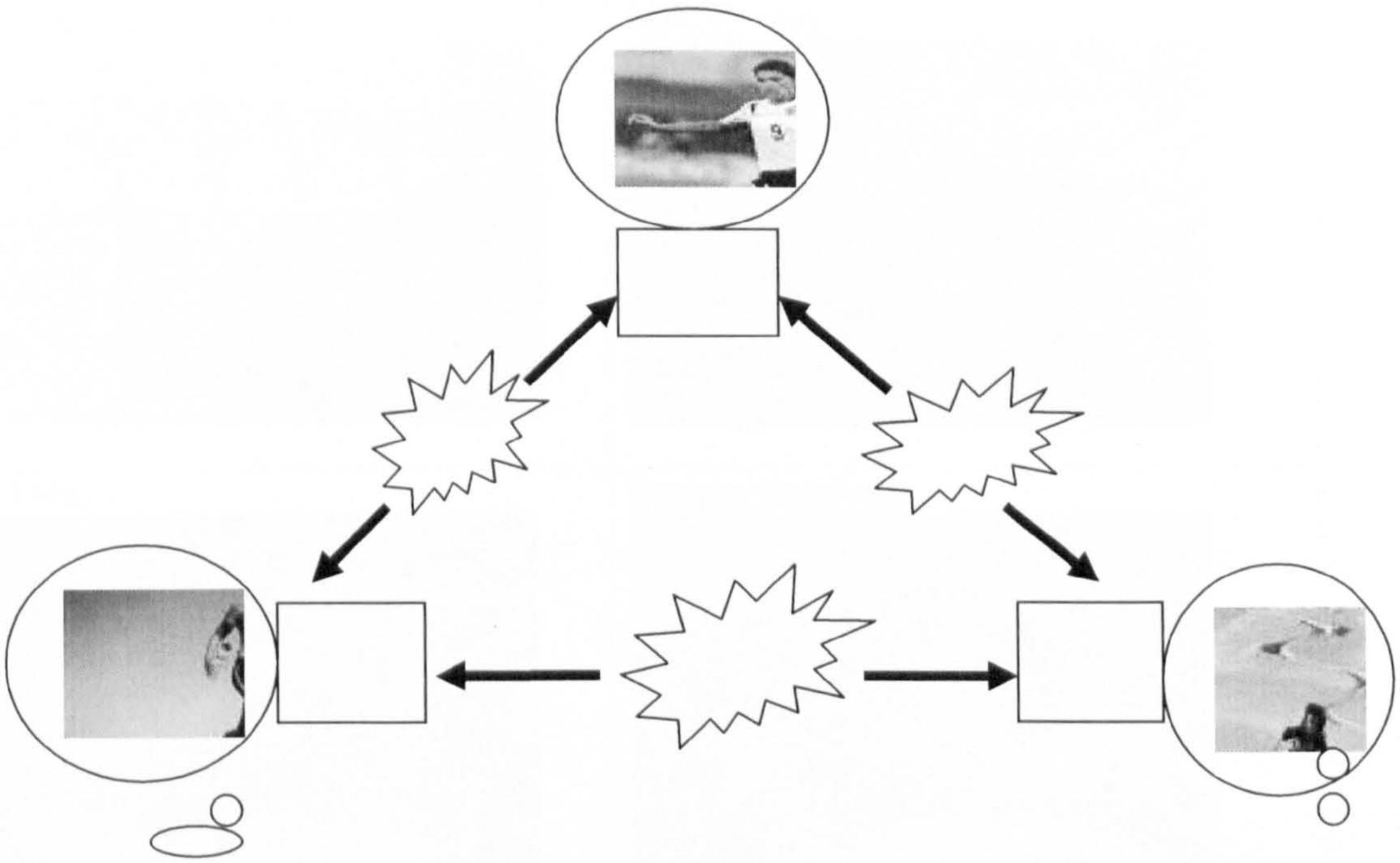
Dogs in the dogs' home

Name:	Fifi 	Fluffy 
Breed:	French poodle	Bloodhound cross
Sex:	Female	Male
Age:	5 years	4 years
Size:	Medium	Large
Diet:	Chicken and ham	Large amounts
Guard dog:	No	Yes
Likes:	Children and other dogs	To sit by the fire
Dislikes:	Rain and the vet	Cats

Name:	Pinky 	Smiley 
Breed:	Beagle	Unknown
Sex:	Female	Female
Age:	10 years	3 years
Size:	Small	Medium
Diet:	Milk and biscuits	Fish and dog food
Guard dog:	No	Yes
Likes:	To be patted	People and playing
Dislikes:	Running and snow	Having to stay in

Name:	Red foot 	Pony tail 
Breed:	Unknown	Unknown
Sex:	Male	Male
Age:	6 months	5 years
Size:	Small	Small
Diet:	Shoes and leftovers	Sausages
Guard dog:	No	Yes
Likes:	Going out to play	Being naughty
Dislikes:	Doing as he's told	Baths

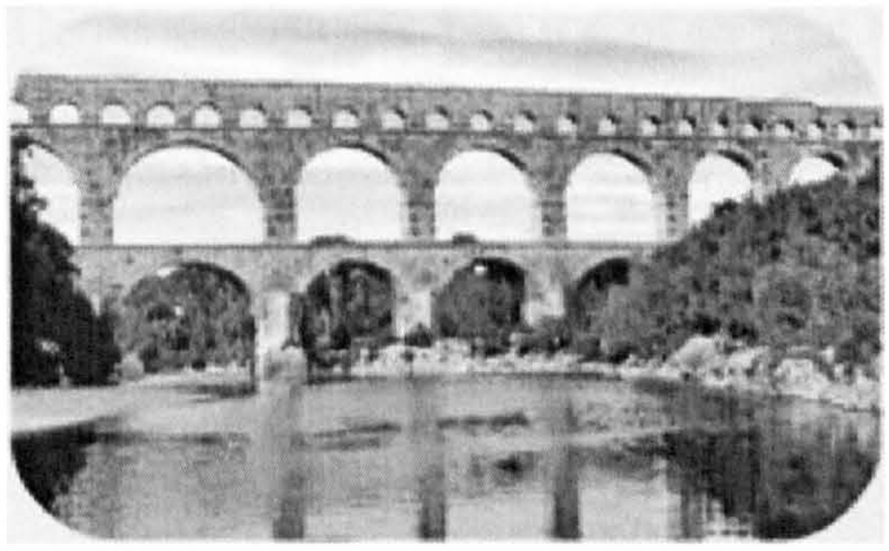
Week 6: Odd One Out – Sports
Thinking skills: analysis (comparing and contrasting)



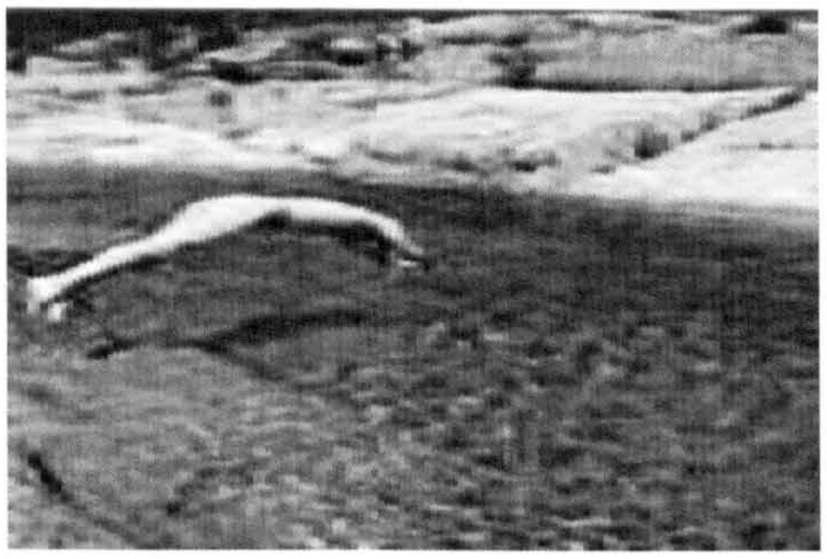
Week 7: Make Up A Story

Task: Use the following 4 pictures to make up your own story in a group.
Thinking skills: synthesis

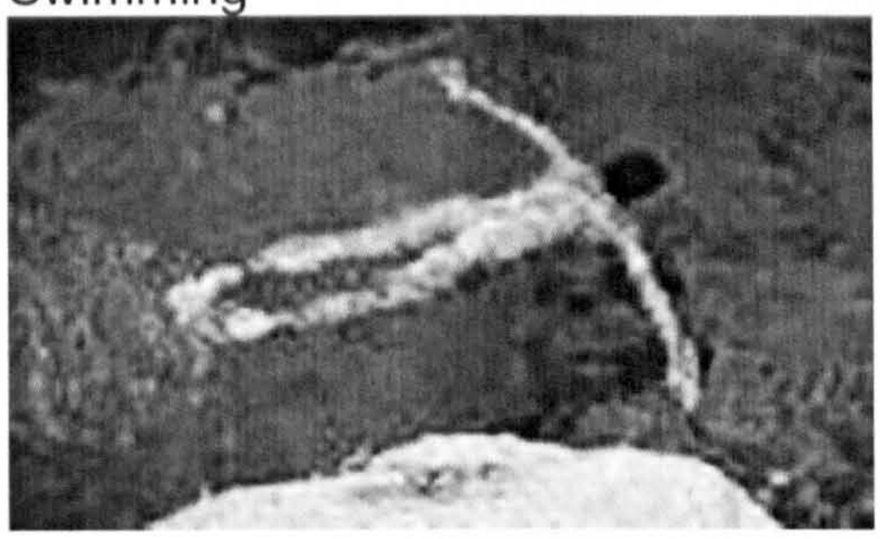
Pont du Gard in France



Naked man



Swimming



People laughing



Vocab.
breaststroke / backstroke/ butterfly stroke/ free style stroke/ doggie style / dive / naked/ swim suit

Sentence pattern / phrases:
I can swim breast stroke / I can do breast stroke
Dive into the water
The bridge has 3 layers.

Task: Based on the pictures provided above, Make Up A Story with 10-20 sentences.

You may start your story as the following or use your own way:
I was standing on the first layer of the bridge, looking….

Week 8: Guess What I Say

Task 1: Guess who/what the celebrity/country/historical site/sport/food is. (Students come to the front and talk about the topic without mentioning the key word, the others guess.)
Thinking skills used by the speaker: comprehension, application
Thinking skills used by the ones who listen and guess: Analysis

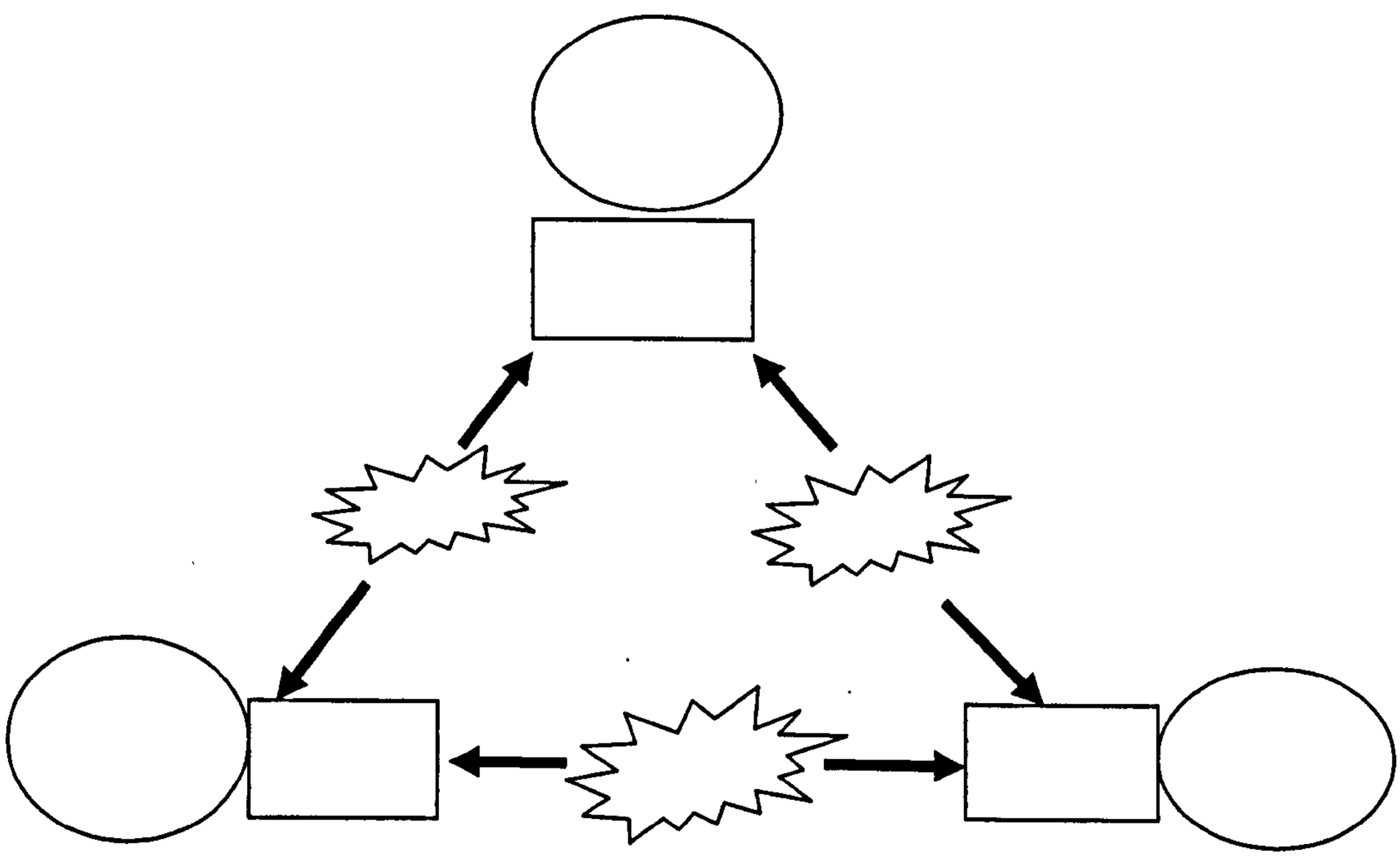
Categories	objects			
Celebrity	王建民 (A baseball player)	Jolin (A singer)	張菲 (TV program host)	David Beckham (A football player)
Country	Canada	Brazil	Singapore	France
Historical site	The Great Wall	Stonehenge	Ankor Watt	Pyramid of the Sun
Sports	Tennis	Skiing	American football	Marathon
Food	Fired rice noodles	Stinky Tofu	Durian	Ginger

Task 2: All students guess what it is by asking questions, e.g. Is it a celebrity/country/historical site/sport? Or Is it a male or female? Etc.
Thinking skills used by the ones who guess: analysis

Categories	objects		
Celebrity	Lin Zi-ling (A model)	A-Mei (A singer)	Liang Hon-Zi (Music producer)
Country	The Philippines	Thailand	India
Historical site	An-Ping castle	Parthenon	Pisa Tower
Sports	Water skiing	Mountain biking	Bungee jumping
Food	Chocolate cake	Don-sun duck head	Wan-Luen pig knuckle

Week 9: Odd One Out

Select your own objects and discuss. Then swap your material with other group and talk about it.

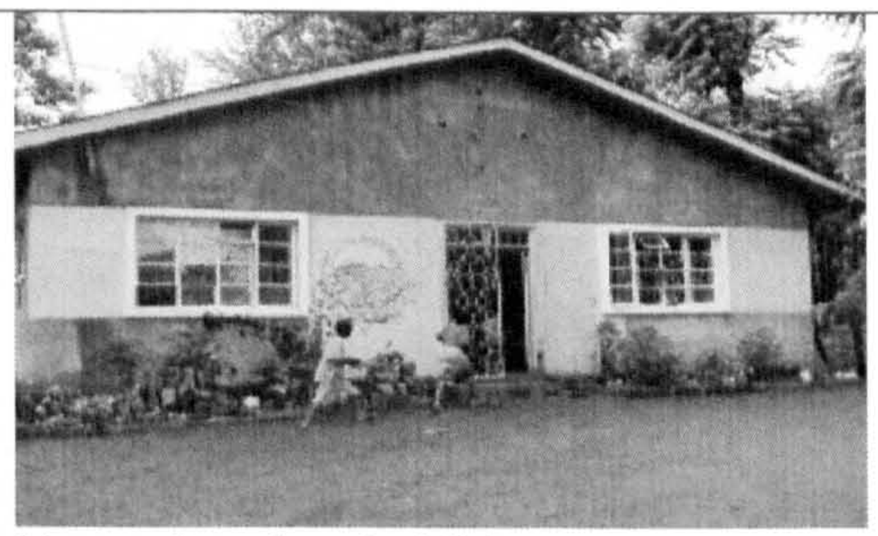


Week 10: 5Ws – Travelling packages

Package 1: **Build homes for street children in Peru**



Street kids



Home for street children



Land of the Inca



Machu Picchu

Who can volunteer?
If you are 18 or over, you can be involved in this exciting project, building homes for street children in Peru.

Activity:
1. Spend two weeks to build homes for children there.
2. Spend 3 days to visit the historical site, Land of Inca and Machu Picchu.

Length: 17 days

Fee: NT$55.800 per person, including flights, accommodation, and sightseeing.

Package 2: Trekking/Hiking in Himalayas





This trip is physically demanding.

Activity:

1. Hiking in Himalayas for 7 days, see magnificent rivers and mountain range

Accommodation:

1. Wild camping x 7 nights,
2. Staying in a hotel x 2 nights
3. Traveling on a night train x 2 nights

Length: 12 days

Location: India

Fee: NT$39,800 per person, including flight, accommodation, and trekking cost

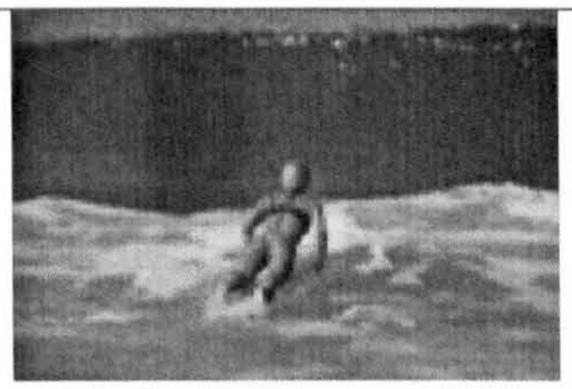
Package 3: Learn to surf


Come to learn surfing


First lesson


practice


Go girl!


Give it a try






Cabin accommodation


Cabin accommodation

Activity:

1. Learn to ride a wave within a day.
2. Surf instructors will progress you through the basics of standing up, paddling, wave selection and ocean knowledge. You will learn to turn, take the drop, and cut across the waves.

Accommodation:

1. Cabin accommodation x 5 nights
2. Beach side camping x 1 night

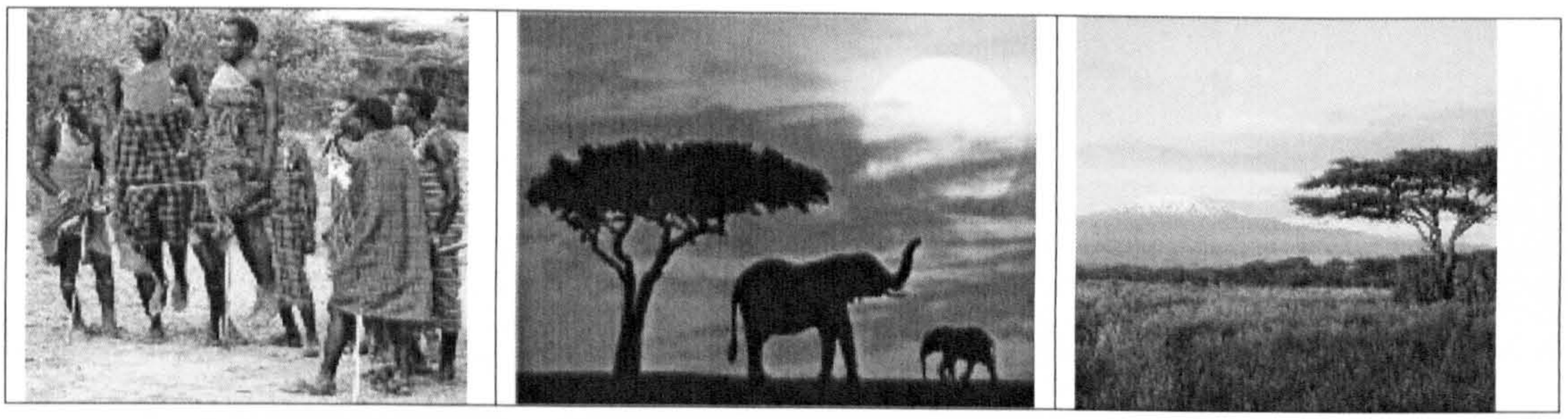
Length: 7 days

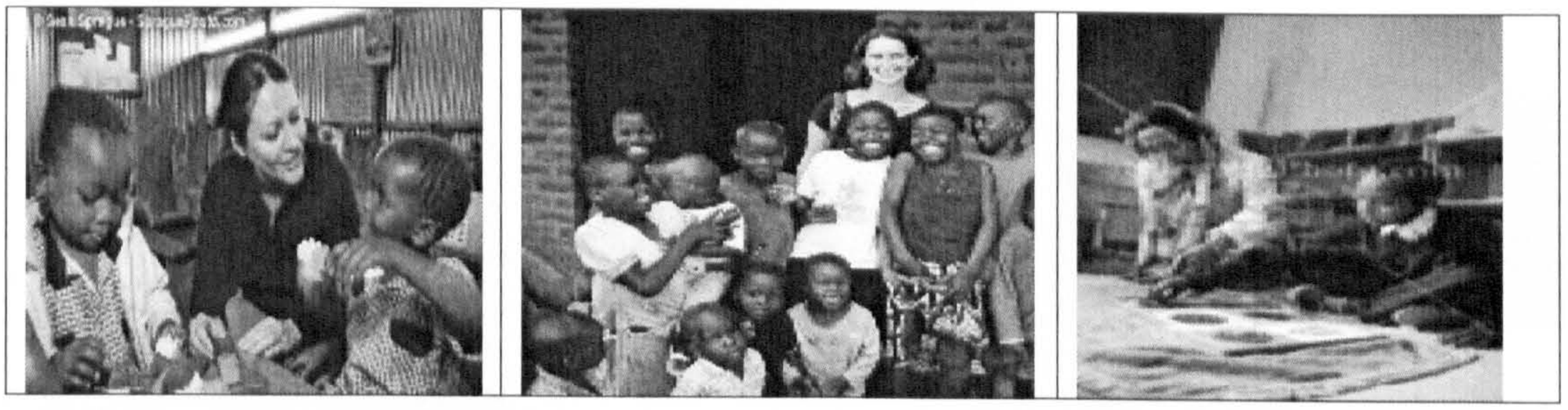
Location: Golden Cost in Australia

Fee: NT$68,200, including flights, accommodation, and surfing lessons

Package 4: Meaningful travel

Project: Community work with children in Kenya





Activity:

1. Help those children with AIDS virus in Kenya.
2. You can provide some basic skills and companionship, e.g. cooking meals, teaching English, art work, playing sports, etc.

You can learn the different culture through this trip and improve the quality of your life by helping others.

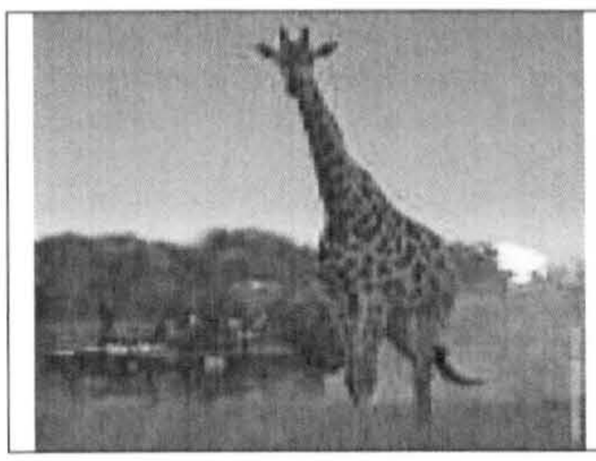
Accommodation: The community provides the accommodation

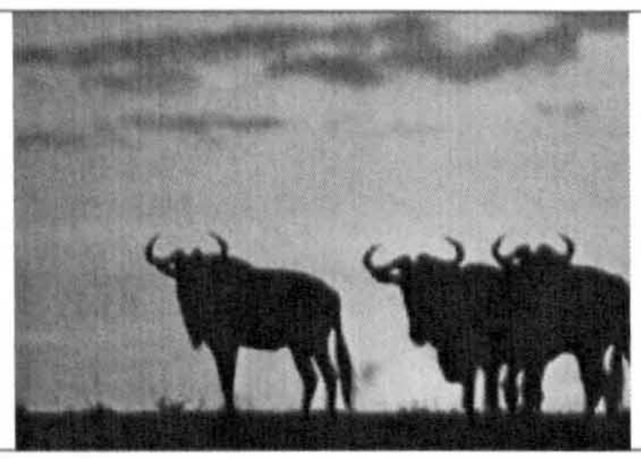
Length: 2 weeks

Fee: NT$34,000 per person, including flights and accommodation

Package 5: Wild animal safari in South Africa

Safari: Riding in a jeep to see wild animals, e.g. rhino, elephant, zebra, buffalo, etc















Safari Lodge

Activity:

1. Riding in a jeep to see wild animals and experience the nature.

Accommodation:

1. Wild camping x 4 nights (no water for shower)
2. Safari Lodge x 2 nights

Length: 7 days

Fee: NT$52,100 per person, including flights, accommodation and meals

Package 6: Sightseeing in Europe

Route: Taipei - Amsterdam (Holland) – Berlin (Germany) – Munich (Germany) – Vienna (Austria) – Geneva (Switzerland) – Nice (France) – Paris (France) –Taipei



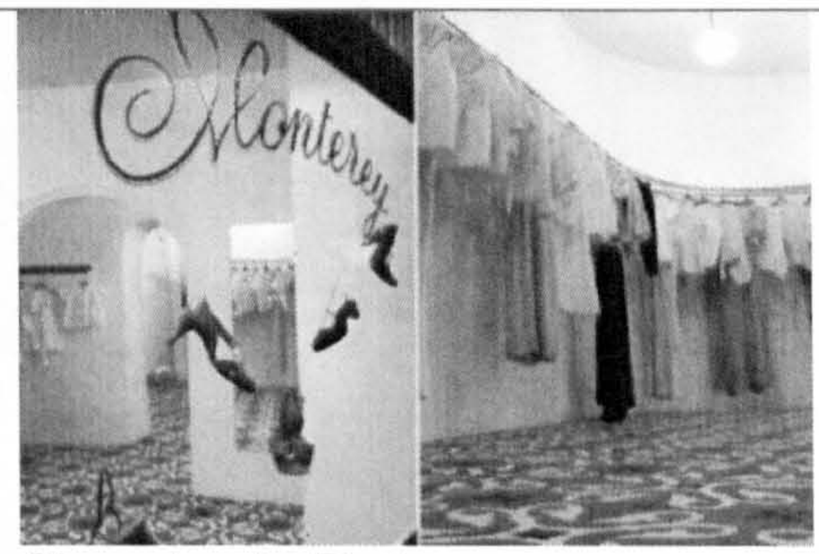
Bicycle-riding in Holland



Jungfrau in Switzerland



Castle in Germany



Shopping in Paris

Activity: bicycle-riding in Holland, cruising on River Rhine, visiting castles, German beer tasting, Opera night in Vienna, visiting Jungfrau (mountain), relaxing at the beach in Nice, shopping in Paris,

Accommodation: 4 start hotels x 13 nights

Length: 15 days

Fee: NT$133,800 per person, including flights, accommodation and meals

Package 7: Summer resort holiday in Fiji

The aim for this package is to get away from your busy life and have relaxation.


Sea world


Resort cabin


Sunbathing


Resort cabin


Local village


Local market

Activity:

1. None (you may do any activities as you wish, such as swimming, sunbathing, diving, visiting local village)
2. Enjoy the sun shine, tropical fruit, local food, and the wonderful sea world

Accommodation: resort cabin x 7 nights

Length: 8 days

Fee: NT$39,500 per person, including flights, accommodation, and breakfast

Task:
Based on the 7 travelling packages provided, choose a suitable package for the role assigned to you. Provide reasons why the package chosen is the most suitable one for the role and why the other packages are not so suitable.

Thinking skills: analysis, evaluation

Role:

1. A middle-aged man, rich and feeling empty inside. He often travels to western countries for sightseeing. He's looking for something which can make his life more exciting.

2. A university student, energetic and sporty. He travels a lot in Taiwan during summer and winter vacation and loves to make friends while travelling. This time he would like to see a different world and experience different culture and life. He's looking for challenge with limited budget.

3. A middle-aged couple with 2 teenage sons. They have a factory, manufacturing clothes. Although they are busy, they often go hiking on the weekends. Their two sons love swimming and hiking. The family would like to experience something different this time.

4. An old couple with a good retirement life. They don't have any children and have kind hearts towards those children without parents. They love travelling and see different culture and life.

5. A young couple with a five-year-old son. They just bought a house last year and need to pay for the mortgage every month. Therefore, they don't have much money left for travelling. They love swimming, trekking, and sightseeing.

6. A fifty-year-old female teacher. She's never been abroad. Experiencing something different apart from her life is what she always wants to do. She's also looking for something which can enrich her life and make her life more meaningful.

Week 11: Guess What I Say

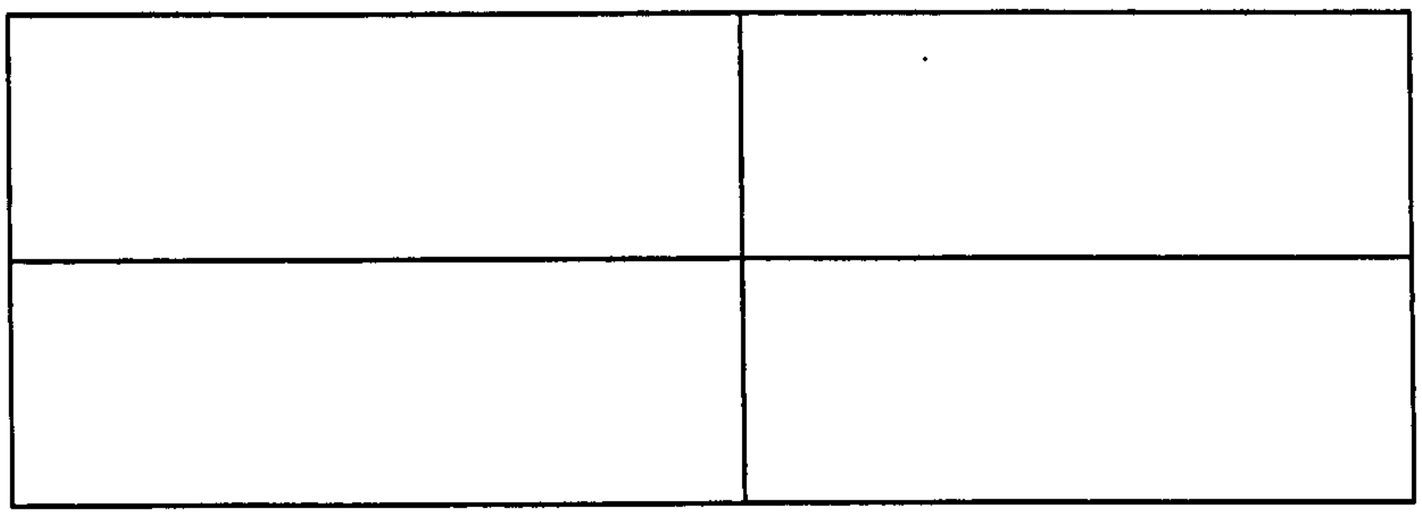
Task: Each student chooses one's own topic, e.g. a person, an object, a place, an idiom, a proverb, food, etc. They take turn to come to the front of the class. One student describes his/her topic without mentioning the key word and the others guess what it is.

Thinking skills used by the speaker: comprehension, application
Thinking skills used by the ones who listen and guess: analysis

Week 12: Make Up A Story

Task: Every student brings one or two any kind of pictures to the class. 4 pictures are provided by individual members in a group, they then make up the story based on these pictures. Following this, pictures are swapped among groups and students make up the story again. After the story is done, two groups get together and share their stories.

Thinking skills: synthesis

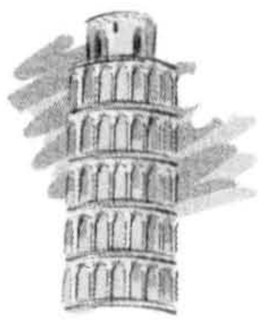


Appendix 2 - An Example of Speaking Activities in the Textbook

Exercise 3 (continued):

5. A gondola in Venice
6. The Golden Gate Bridge
7. The Taj Mahal
8. Stonehenge
9. The Statue of Liberty
10. Mount Rushmore
11. The Leaning Tower of Pisa
12. The Arc de Triomphe
13. St. Basil's Cathedral
14. The Parthenon
15. The Pyramids
16. The Sphinx
17. Tower Bridge
18. Bullfighting
19. David


e.

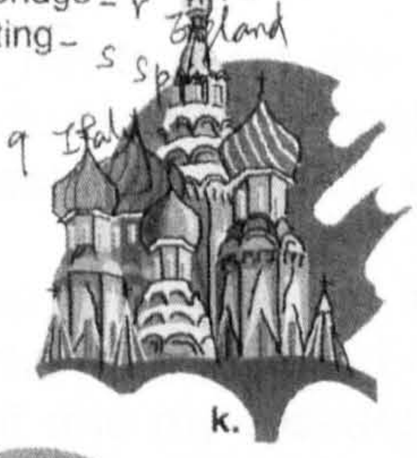

f.


g.

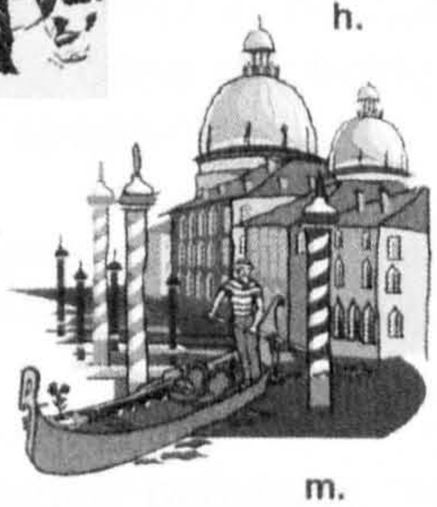

h.

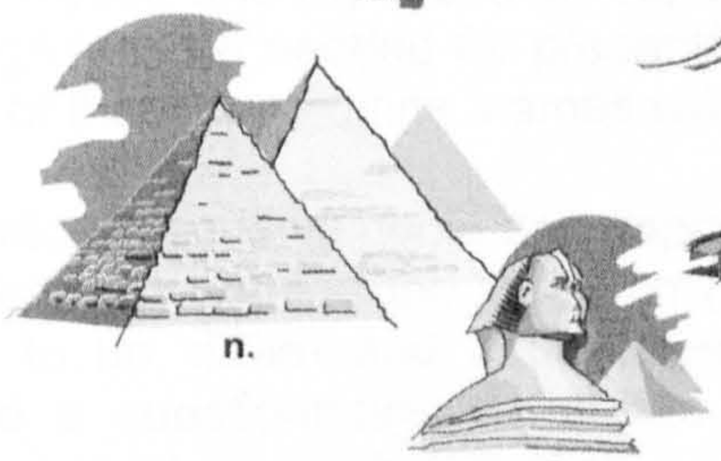

j.


i.

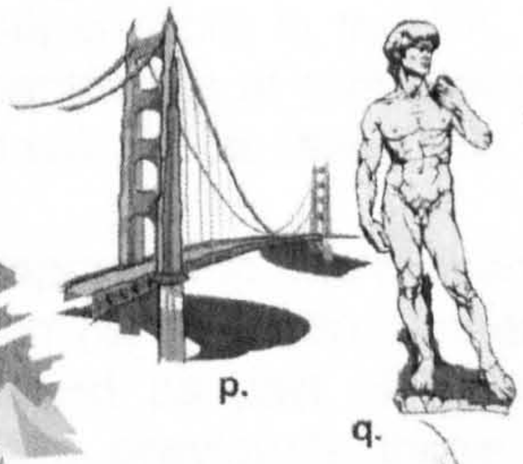

k.


l.

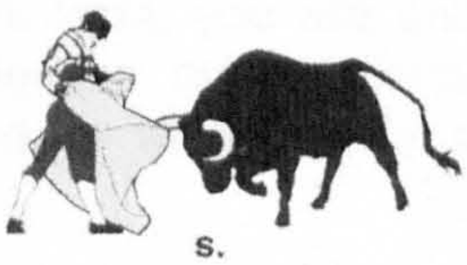

m.


n.

o.


p.


r.


s.

7 Discussion

1. Discuss with your partner in (or near) which cities you can see some of these monuments or sights.
2. Which monuments have you seen or visited in real life? Which ones would you like to visit? Why?
3. Which countries do you think are: a. the three most interesting to visit?
   b. the three cheapest to visit?
   c. the three most dangerous to visit?
   d. the three with the most interesting historical sights?

8 Assignment

Choose one of these monuments and find out as much as you can about it (and maybe the country it comes from) in a book or over the Internet. Next class, form groups and tell your classmates what you found out (your classmates should ask you questions about it, too). Your teacher will then choose a few students to present their information to the class.

Appendix 3 - Letter of Consent and Consent Form

Letter of Consent for the Comparison Class

Newcastle University

Dear students,

My name is Mei-Hui Chen. I am currently studying for an IPhD in Educational and Applied Linguistics at Newcastle University in the UK. Your class are being invited to take part in my study investigating L2 (Second Language) speaking.

This study will not take up any extra time unless you agree to be interviewed. I will video- and audio-tape your English class three times while you are having a lesson, distribute a questionnaire twice, and interview volunteers.

These data will be used only for educational and research purposes, and will be stored in a locked cabinet and destroyed in two years. In the main they will only be seen by myself and my research supervisors in the UK. Sometimes, excerpts of the recordings may be needed for presentations at professional conference. However, in any use of these recordings, names will not be identified.

Your participation is valuable and important to this research. You are free to withdraw at any time. By signing the research participation consent form, you are giving your consent to be video- and audio-taped as part of the lesson observations and to complete a questionnaire. As stated previously these data will be only used for educational and research purposes.

Your kind cooperation will contribute to the English language teaching of university level in Taiwan. It is greatly appreciated.

Mei-Hui Chen
IPhD in Educational and Applied Linguistics
School of Education, Communication and Language Sciences
TEL: 0191 222 5672
E-mail: Mei-Hui.Chen@ncl.ac.uk

Newcastle University

說明書

親愛的同學們，

我是英國新堡大學博士班教育及應用語言學研究生，陳美惠。 貴班被應邀參與這項研究，這項研究旨在探討英文'說'的技巧。

此研究不會佔用您額外的時間， 除非您同意接受訪談。 收集資料時間為一學期。 在此期間將錄影三次，做問卷調查兩次，必要時會訪談自願參與的同學。

這些資料只用在教育研究目的，我會妥當保存並在兩年之後銷毀。 這些資料只供我和我英國的指導教授研究閱讀。 有時資料可能會再專業研討會上公開，但您的名字絕對是被保密的。

您的參與對這項研究是非常有價值及重要性，您可以隨時退出參與此研究。 簽此同意書表示您願意接受錄影及訪問， 和將這些資料用在教育研究目的之專業研討會上。 如先前提到的，這些資料只用在教育研究目的。

您的參與將對台灣之大學英語教育有所貢獻。 感謝您的合作。

陳美惠
英國新堡大學，教育溝通及語言科學系所，教育及應用語言學組
電話: 0191 222 5672
E-mail: Mei-Hui.Chen@ncl.ac.uk

Letter of Consent for the Innovation Class

Dear students,

My name is Mei-Hui Chen. I am currently studying for an IPhD in Educational and Applied Linguistics at Newcastle University in the UK. Your class are being invited to take part in my study investigating L2 (Second Language) speaking.

Your lecturer, Chuen-Hui Tai, and I are cooperating in conducting this study. The main aim is to explore how thinking skills can enhance L2 speaking by using HOTS approach. HOTS approach provides opportunities for you to use higher-order thinking and express your opinions and ideas, and this will enable you to use the language meaningfully.

This study will be conducted for one semester with the intervention of HOTS approach into your timetabled lessons. This will not take up any extra time unless you agree to be interviewed. I will video- and audio-tape your English class three times while you are having a lesson, distribute a questionnaire twice, and interview volunteers.

These data will be used only for educational and research purposes, and will be stored in a locked cabinet and destroyed in two years. In the main they will only be seen by myself and my research supervisors in the UK. Sometimes, excerpts of the recordings may be needed for presentations at professional conference. However, in any use of these recordings, names will not be identified.

Your participation is valuable and important to this research. You are free to withdraw at any time. By signing the research participation consent form, you are giving your consent to be video- and audio-taped as part of the lesson observations and to complete a questionnaire. As stated previously these data will be only used for educational and research purposes.

Your kind cooperation will contribute to the English language teaching of university level in Taiwan. It is greatly appreciated.

Mei-Hui Chen
IPhD in Educational and Applied Linguistics
School of Education, Communication and Language Sciences
TEL: 0191 222 5672
E-mail: Mei-Hui.Chen@ncl.ac.uk

Newcastle University

說明書

親愛的同學們，

我是英國新堡大學博士班教育及應用語言學研究生， 陳美惠。 貴班被應邀參與此研究，這項研究旨在探討英文'說'的技巧。

此研究將由戴春惠老師和我一起合作實行。 這研究主要用'高層思考詢問'來探討'思考技巧'如何促進'英文說的能力'。 '高層思考詢問'提供同學們使用高層思考及表達意見的機會，這並且能使您有意義地使用語言。

實驗時間為一學期，這將不會佔用同學額外的時間，除非您同意參與訪談。 在此期間將錄影三次，做問卷調查兩次，並訪談自願參與的同學。

這些資料只用在教育研究目的，我會妥當保存並在兩年之後銷毀。這些資料只供我和我英國的指導教授研究閱讀。 有時資料可能會再專業研討會上公開，但您的名字絕對是被保密的。

您的參與對這項研究是非常有價值及重要性，您可以隨時退出參與此研究。 簽此同意書表示您願意接受錄影及訪問， 和將這些資料用在教育研究目的之專業研討會上。 如先前提到的，這些資料只用在教育研究目的。

您的參與將對台灣之大學英語教育有所貢獻。 感謝您的合作。

陳美惠
英國新堡大學，教育溝通及語言科學系所，教育及應用語言學組
電話: 0191 222 5672
E-mail:Mei-Hui.Chen@ncl.ac.uk

Newcastle University

Research Participation Consent Form
Infusing thinking skills into an L2 Classroom

Please read through the following statements and tick in the box.

1. I have read the description and understand the information given which describes what this study is about and data collection methods will be taken. ..... ☐

2. My participation is voluntary and I feel free to withdraw at any time..... ☐

3. I agree to take part and cooperate in this study ..... ☐

4. I agree to be video- and audio-taped of my voice and face and to be interviewed. ..... ☐

_____	_____	_____
(Name)	(Signature)	(Date)

Newcastle University

參與研究之同意書

思考技巧與高層思考詢問在外語教室的影響

請閱讀下列條文並在空格裡打勾

1. 我已閱讀並了解此研究之目的，及資料收集之方式..... □

2. 我的參與是自願性的，並可隨時退出..... □

3. 我同意參與並全力配合此研究..... □

4. 我同意配合錄影及錄音， 並接受訪談..... □

_____ _____ _____

(姓名) (簽名) (日期)

Appendix 4 – Thinking Tasks Used for Data Collection

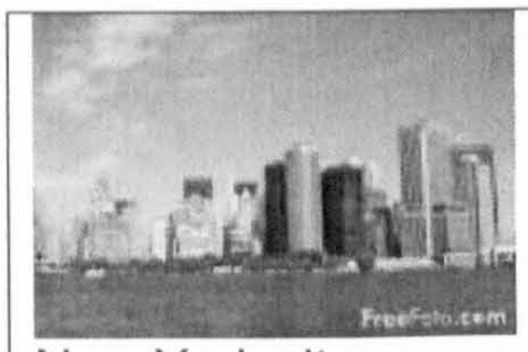
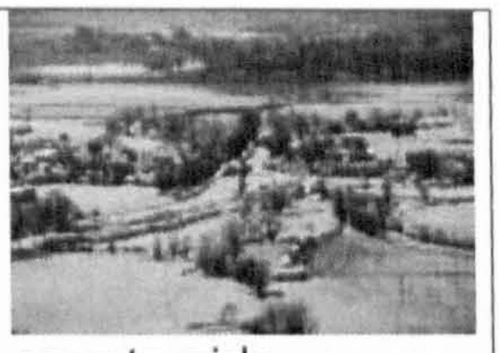
Pre-test Task

Topic: The most ideal country you would like to live in

Thinking skills: analysis, evaluation

Criteria:

1. low living cost (cheap alcohol, cigarette, medication, food)
2. low suicide rate
3. accept women to have a child without getting married
4. mountains
5. forest
6. sand beaches
7. hot weather
8. snowing
9. famous universities
10. safety
11. attractive men / women
12. low murder rate
13. good, cheap public transportation
14. good salary
15. more opportunities for jobs
16. good social welfare
17. free education
18. stable democracy
19. communism
20. famous for fashion
21. famous for food
22. famous for country life
23. natives are friendly to foreigners
24. fantastic social life/ night life

 New York city	 Beach resorts	 Cherry bloom	 countryside
 Canadian Mountain cabin	 Snow view	 Peru – Andes	 Desert

Among the countries you are familiar with, choose a country you would like to live in beyond Taiwan. You may consider the factors, like, environment, political situation, economic situation, education, social life, safety, transportation, weather, food, etc.

Task:

1. Create a most ideal country you would like to live in based on 6 most important criteria you choose. Provide with reasons of why these 6 criteria are the most important.

2. Which country in this world is the most similar to the ideal country you create? Provide evidence.

3. What are the similarities and differences between Taiwan and the country you choose to live in?

Post-test task

Topic: Personalities
Thinking skills: evaluation, analysis

* fussy * possessive * active * passive * bossy (assertive) * funny * boring * hard-working * lazy * intelligent * friendly * unfriendly * kind * unkind * macho	● outgoing ● easy-going ● reserved ● greedy ● generous ● stingy ● caring ● supportive ● loyal / faithful ● disloyal / unfaithful ● reliable ● unreliable ● understanding/ sympathetic ● like teasing / laughing at others ● aggressive (easily attack others; work hard to be successful)

Discuss the following questions:

1. Choose 3 most important characters you think that a good friend should have. Choose 1 character you think that a good friend shouldn't have. Provide reasons. (You need to reach an agreement within your own group.)

2. Choose 3 most important characters you think that a good husband/wife should have. Choose 1 character you think that a good husband/wife shouldn't have. Provide reasons. (You need to reach an agreement within your own group).

3. Can you explain why the characters you choose for being a good friend and partner are different?

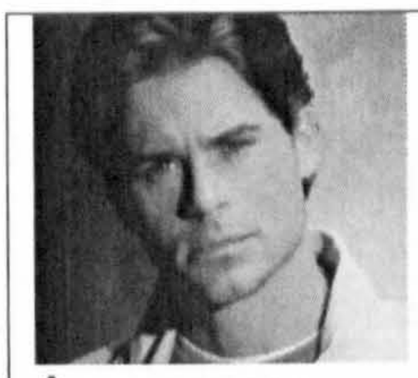
Delayed Post-test Task

Topic: An ideal mate
Thinking skills: Evaluation, analysis

Question:
1. Suppose you are a female (girl/woman), think about what criteria are the most important for you when choosing an ideal mate. In your group, choose 3 most important criteria. You might have different opinion from your group members. You need to persuade them and to reach an agreement in your group.
2. Talk about the reasons why the other criteria are less important or not important.

The following are some criteria when looking for an ideal mate.

- appearance: good looking, handsome, macho,
- character: responsible, caring, understanding, funny, high EQ
- high education: someone with a master or phd degree
- family background: someone from a rich family or poor farmer's family, big family
- occupation: lawyer, doctor, engineer, etc
- hobbies: mountain climbing, travelling, etc.
- health condition:
- soulmate
- nationality: Taiwanese, foreigner
- others


Appearance


Character


High education


Family background


Occupation


hobbies


Health condition


Financial status

知己
soulmate



Nationality

Appendix 5 - Questionnaires

Pre-test Questionnaire Gender:

This questionnaire is to investigate your attitudes and beliefs towards thinking skills and questioning-answering behavior. It is anonymous. There is no right or wrong answer, your **true answer** is the best.

This questionnaire contains three sections:
Section 1: Related to your beliefs and attitudes towards thinking skills used in English class.
Section 2: Related to what type of questions you can answer in English class.
Section 3: Related to your attitudes towards questioning-answering behavior in English class.

Section 1

This section asks **your attitudes towards thinking skills you use in English class.** There are six scales to choose from, 1= Strongly agree, 2= Agree, 3= Slightly agree, 4= Slightly disagree, 5= Disagree, 6= Strongly disagree. Please choose an appropriate answer and tick (v) in the box.

Questions \ Scales	1 Strongly agree	2 Agree	3 Slightly agree	4 Slightly disagree	5 Disagree	6 Strongly disagree
1. It is important to **memorize what has learned in English class.**						
2. It is important to **recite** English passages to practice English.						
3. It is important to be able to **recall what I have learned in** English class.						
4. It is important to **understand** the learning materials in English class.						
5. Being able to **summarize** what we have read or listened to is important in English class.						
6. Being able to **retell** what we have read or listened to is important in English class.						
7. It is important to **apply** what we have learned in English class, e.g. to use the reading strategies learned in class to read English newspaper after class.						
8. Being able to **think what I would do** as a character in a novel is important in English class.						
9. It is important to be able to think 'alternative ways' to explain when people do not understand me in English class.						
10. It is important to be able to give **reasons** in English class, e.g. to justify a decision I have made.						
11. It is important to use **comparison and contrast** skills in English class, e.g. to tell similarities and differences of the two characters in a novel.						
12. It is important to be able to **analyze** articles in English class, e.g. to analyze the strengths and weaknesses of a plan.						

Questions \ Scales	1 Strongly agree	2 Agree	3 Slightly agree	4 Slightly disagree	5 Disagree	6 Strongly disagree
13. Being able to **solve problems** is important in English class, e.g. to think about a solution to save money.						
14. **Creativity** is important is English class, e.g. to make a new ending for a story.						
15. It is important to be able to **design** things in English class, e.g. to design an itinerary for a two-week holiday.						
16. Being able to **assess** is important in English class, e.g. to assess which travelling package is better for students and provide reasons and evidence.						
17. It is important to be able to make a good **argument supported** with reasons or evidence, e.g. to argue for not having dress code on campus and provide with reasons or evidence.						
18. It is important to be able to make **judgments** based on evidence or reasons, e.g. to show my position, agree or disagree, and give reasons or evidence.						

Section 2

This section asks about **different types of questions you can answer in English class**. There are six scales to choose from,1= Strongly agree, 2= Agree, 3= Slightly agree, 4= Slightly disagree, 5= Disagree, 6= Strongly disagree. Please choose an appropriate answer and tick (v) in the box.

Questions \ Scales	1 Strongly agree	2 Agree	3 Slightly agree	4 Slightly disagree	5 Disagree	6 Strongly disagree
1. I **can answer** questions which I have memorized the answer in English class, e.g. 'Who inherited a fortune from Mr. Brown in this story?'						
2. I **can answer** questions which ask for information I have memorized, e.g. 'Recite the text.'						
3. I **can answer** questions where the answer can be found in the text, e.g. 'What did the old man say before he died in the text?'						
4. I **can answer** questions which ask for understanding of the materials learned in English class, e.g. 'What does this sentence mean?'						
5. I **can answer** questions which require a summary of an article or a story read or listened to in English class, e.g. "What is the main idea of this article?'.						

Questions \ Scales	1 Strongly agree	2 Agree	3 Slightly agree	4 Slightly disagree	5 Disagree	6 Strongly disagree
6. I **can answer** questions where I need to retell what I learned in English class, e.g. 'Describe what happened to World Trade Center in New York on September 11, 2001.'						
7. I **can answer** questions where I need to apply things I have learned to a new context in English class, e.g. 'How is O zone layer related to global warming?'						
8. I **can answer** questions where I need to place myself in other situation, e.g. 'What would you do, if you were this person?'						
9. I **can answer** questions where I need to provide a new example in English class, e.g. 'Can you give me another example?'						
10. I **can answer** questions where I need to give reasons in English, e.g. 'Why do you make this decision?'						
11. I **can answer** questions which require comparison and contrast skills in English class, e.g. 'Can you tell the similarities and differences between these two characters in this story?'						
12. I **can respond to** questions where I need to analyze in English class, e.g. 'What are the strengths and weakness of merging these two companies?'						
13. In English class I **can respond to** questions where I need to think of a possible solution to a problem, e.g. 'How can you improve your English speaking ability?'						
14. I **can answer** questions which require creative skills, e.g. 'Can you make up a new ending for this English novel?'						
15. I **can respond to** questions where I need to design something new in English class, e.g. 'How can you design another new itinerary based on the budget we have?'						
16. I **can answer** questions which require assessment and reasoning skill, e.g. 'Which do you think is better and why?'						
17. I **can answer** questions where arguing ability is required and I need to elaborate my views and provide reasons or evidence in English class, e.g. 'What is your opinion towards this plan? Provide reasons or evidence.'						
18. I **can respond to** questions where I need to give my personal views, e.g. agree or disagree, and give reasons to support my views. e.g. 'Do you agree with this decision? Why?'						

Section 3

This section asks about **your belief about questioning-answering behavior in English class**. There are six scales to choose from, 1= Strongly agree, 2= Agree, 3= Slightly agree, 4= Slightly disagree, 5= Disagree, 6= Strongly disagree. Please choose an appropriate answer and tick it in the box.

Questions \ Scales	1 Strongly agree	2 Agree	3 Slightly agree	4 Slightly disagree	5 Disagree	6 Strongly disagree
1. I can **talk freely in front of the class** when answering questions where the answers **can** be easily found in the text in English class.						
2. I can **express** my own opinions and ideas **freely in front of the class** when answering questions where the answers to the questions **can not** be easily found in the text in English class.						
3. In English class when the teacher asks a question where the answer **can not** be found in the text, the teacher usually gives us **enough time to think** about the question and answer it.						
4. After a **student expressed his/her opinions**, the teacher usually gives us **enough time to think** about what this student has said and allows us to comment or elaborate our own opinions in English class.						
5. After one student expressed his/her opinions, the teacher usually **asks** him/her **for additional information** in English class.						
6. 'Questions' which are **intellectually challenging** are interesting.						
7. I like to answer questions in English which are intellectually challenging.						
8. 'Questions' which are **intellectually challenging** provide me **more opportunities to talk in English.**						
9. 'Questions' which are intellectually challenging are useful in terms of **improving students' thinking** ability, e.g. analysis, creation, and evaluation.						

前測問卷調查　　**性別:**

本問卷旨在探討您對思考技巧及問與答的態度和意見，此為不記名問卷，並無對錯之分，您**最真實的回答**，就是最好的答案。

這份問卷包含三部份:
第一部份是關於在英文課堂中，您對所使用的思考技巧的看法及態度。
第二部份是關於在英文課堂中，您回答不同類型問題的能力。
第三部份是關於在英文課堂中，您對問與答的看法。

第一部份

第一部份是關於在「**英文課堂**」中，**您對所使用的思考技巧的看法及態度。** 共六個選項可供選擇（1=非常同意，2=同意，3=少許同意，4=少許不同意，5=不同意，6=非常不同意），請選取一個合適的答案，並在格子裡打勾(v)。

問題 \ 程度	1 非常同意	2 同意	3 有點同意	4 有點反對	5 反對	6 非常反對
1 英文課中，用「**記憶力**」記住所學的東西，是很重要的。						
2 英文課中，使用「**背誦能力**」是很重要的。如，背誦英文課文。						
3 使用「**回想能力**」，回想起英文課中所學的東西是很重要的。						
4 英文課中，能用「**理解力**」理解文章或教材的意思，是很重要的。						
5 英文課中，能用「**做摘要的能力**」將所閱讀或聽到的英文文章簡短地說出其概要, 是很重要的。						
6 英文課中，能用「**敘述的能力**」將所閱讀或聽到的文章用英文重述一遍，是很重要的。						
7 英文課中，會使用「**應用所學的能力**」是很重要的。如，將所學到的閱讀技巧，應用在課後閱讀英文文章上。						
8 英文課中，能想像自己若是小說中的主角，會怎麼做，這種使用「**設身處地思考的能力**」，是很重要的。						
9 英文課中，當同學不了解我的意思時，我能用「**其他例子再解釋一次**」，這種能力是很重要的。						
10 英文課中，能「**說明原因及理由的能力**」是重要的。 例, 當我做一項決定時，我能說出做此決定的原因。						
11 英文課中，能使用「**比較的能力**」將兩件事做比較，並指出其異同點，是很重要的。 如，比較小說中兩角色間的差異。						
12 英文課中，能用「**分析的能力**」分析文章，是很重要的。如，分析文章中一項計劃的優缺點。						
13 英文課中，能使用「**解決問題的能力**」來解決問題，是很重要的。如，想出省錢的方法。						
14 英文課中，使用「**創造力**」是很重要的。如，重新改寫故事結局。						
15 英文課中，使用「**設計的能力**」是很重要的。如，設計兩週假期的旅行行程。						

問題 \ 程度	1 非常同意	2 同意	3 有點同意	4 有點反對	5 反對	6 非常反對
16 英文課中，會使用「**評估能力**」並提出理由及証據是很重要的。如，指出哪個旅行方案較適合學生，並提出理由及証據。						
17 英文課中，能使用「**辯論能力**」並提出理由及証據是很重要的。如，辯論校園中不該有服裝規定，闡述自己的意見，並針對其意見，提供相關的理由及証據。						
18 英文課中，會用「**判斷能力**」是很重要的。如，表達自己的立場，同意或不同意，並針對其立場，提供相關的理由及証據。						

第二部份

第二部份是關於在「**英文課堂**」中，**您回答不同類型問題的能力**，有六個選項可供選擇（1=非常同意，2=同意，3=少許同意，4=少許不同意，5=不同意，6=非常不同意）。請選取一個合適的答案，並在格子裡打勾(v)。

問題 \ 程度	1 非常同意	2 同意	3 有點同意	4 有點反對	5 反對	6 非常反對
1 英文課中，我**能用英文回答**有關我已記憶的東西。如，'故事中，誰從布朗先生那裡繼承了一大筆財富?'						
2 英文課中，我**能用英文回答**有關我已背誦的東西。如，'背出課文'。						
3 英文課中，我**能用英文回答**其答案是在課文裡的問題。如，'課文中，那老人死之前說了什麼?'						
4 英文課中，我**能用英文回答**需要解釋課文內容的問題。如，'這句子是什麼意思?'						
5 英文課中，我**能用英文回答**需做摘要的問題。如，'這篇文章的大意是什麼?'						
6 英文課中，我**能用英文回答**需敘述我所學的東西的問題。如，'描述 2001 年 9 月 11 日紐約世貿中心發生了什麼事。'						
7 英文課中，我**能用英文回答**需將所學運用到新的情境的問題。如，'臭氧層與全球暖化有什麼關係?'						
8 英文課中，我**能用英文回答**需假設自己在另一情境中的問題。如，'若你是這個人，你會怎麼做?'						
9 英文課中，我**能用英文回答**需提供新例子的問題。如，'你可以舉其他例子嗎?'						
10 英文課中，我**能用英文回答**需提供原因及理由的問題。如，'你為什麼要做這個決定?'						
11 英文課中，我**能用英文回答**需做比較的問題。如，'你可以比較故事中這兩個主角個性的異同點嗎?'						
12 英文課中，我**能用英文回答**需做分析的問題。如，'這兩家公司合併，優缺點為何?'						

問題 \ 程度	1 非常同意	2 同意	3 有點同意	4 有點反對	5 反對	6 非常反對
13 英文課中，我**能用英文回答**需思考解決方案的問題。如，'你如何改善英文口語能力?'						
14 英文課中，我**能用英文回答**需用創造力來思考的問題。如，'你能為這英文小說編寫一新結局嗎?'						
15 英文課中，我**能用英文回答**需用設計力來思考的問題。如，'依現有的旅行預算，你如何設計另一新的旅行路線?'						
16 英文課中，我**能用英文回答**需用評估能力的問題，辨別哪個較好，並提出理由及証據。如，'你認為哪個較好? 為什麼?'						
17 英文課中，我**能用英文回答**需用辯論能力的問題，闡述我的意見，並針對其意見，提供相關的理由及証據。如，'你對這計劃有何看法? 並提出理由及証據。'						
18 英文課中，我**能用英文回答**需用判斷能力的問題，表達我的立場，同意或不同意，並針對其立場，提供相關的理由及証據。如，'你同意這決定嗎? 為什麼?'						

第三部份

第三部份是關於在「**英文課堂**」中，**您對問與答的看法**。 有六個選項可供選擇（1=非常同意，2=同意，3=少許同意，4=少許不同意，5=不同意，6=非常不同意），請選取一個合適的答案，並在格子裡打勾(V)。

問題 \ 程度	1 非常同意	2 同意	3 有點同意	4 有點反對	5 反對	6 非常反對
1 英文課中，我可以**自在地在班上用英文回答**課本上**有答案**的問題。						
2 英文課中，我可以**自在地在班上用英文回答**課本上**沒有答案**的問題，這種答案是需自己想出來的。						
3 英文課中，當老師問的問題，其答案是**課本上找不到的**，老師通常給我們**足夠時間思考**，然後再回答。						
4 英文課中，在**學生表達其看法之後**，老師通常給我們**足夠時間思考**那位同學的說法，並讓我們去評論或表達其看法。						
5 英文課中，在學生表達其看法之後，老師通常會**詢問更進一步的相關訊息**。						
6 具**腦力挑戰**的問題，其答案是須靠自己動腦筋想出來的，這種問題是很有趣的。						
7 我**喜歡用英文回答**具腦力挑戰或思考的問題。						
8 具**腦力挑戰**的問題，提供我更多**用英文表達說話的機會**。						
9 具腦力挑戰的問題，在**改善學生思考技巧**很有幫助。如，改善運用、分析、創造、及評論等思考技巧。						

Post-test Questionnaire Gender:

This questionnaire is to investigate your attitudes and beliefs towards thinking skills, questioning-answering behaviour, and the intervention conducted. It is anonymous. There is no right or wrong answer, your **true answer** is the best.

This questionnaire contains four sections:
Section 1: Related to your beliefs and attitudes towards thinking skills used in English class.
Section 2: Related to what type of questions you can answer in English class.
Section 3: Related to your attitudes towards questioning-answering behaviour in English class.
Section 4: Related to your opinions towards the intervention conducted.

Section 1

This section asks **your attitudes towards thinking skills you use in English class.** There are six scales to choose from, 1= Strongly agree, 2= Agree, 3= Slightly agree, 4= Slightly disagree, 5= Disagree, 6= Strongly disagree. Please choose an appropriate answer and tick (v) in the box.

Questions \ Scales	1 Strongly agree	2 Agree	3 Slightly agree	4 Slightly disagree	5 Disagree	6 Strongly disagree
1. It is important to **memorize what has learned in English class.**						
2. It is important to **recite** English passages to practise English.						
3. It is important to be able to **recall what I have learned in** English class.						
4. It is important to **understand** the learning materials in English class.						
5. Being able to **summarize** what we have read or listened to is important in English class.						
6. Being able to **retell** what we have read or listened to is important in English class.						
7. It is important to **apply** what we have learned in English class, e.g. to use the reading strategies learned in class to read English newspaper after class.						
8. Being able to **think what I would do** as a character in a novel is important in English class.						
9. It is important to be able to think 'alternative ways' to explain when people do not understand me in English class.						
10. It is important to be able to give **reasons** in English class, e.g. to justify a decision I have made.						
11. It is important to use **comparison and contrast** skills in English class, e.g. to tell similarities and differences of the two characters in a novel.						
12. It is important to be able to **analyse** articles in English class, e.g. to analyse the strengths and weaknesses of a plan.						

Scales / Questions	1 Strongly agree	2 Agree	3 Slightly agree	4 Slightly disagree	5 Disagree	6 Strongly disagree
13. Being able to **solve problems** is important in English class, e.g. to think about a solution to save money.						
14. **Creativity** is important is English class, e.g. to make a new ending for a story.						
15. It is important to be able to **design** things in English class, e.g. to design an itinerary for a two-week holiday.						
16. Being able to **assess** is important in English class, e.g. to assess which travelling package is better for students and provide reasons and evidence.						
17. It is important to be able to make a good **argument supported** with reasons or evidence, e.g. to argue for not having dress code on campus and provide with reasons or evidence.						
18. It is important to be able to make **judgement** based on evidence or reasons, e.g. to show my position, agree or disagree, and give reasons or evidence.						

Section 2

This section asks about **different types of questions you can answer in English class.** There are six scales to choose from,1= Strongly agree, 2= Agree, 3= Slightly agree, 4= Slightly disagree, 5= Disagree, 6= Strongly disagree. Please choose an appropriate answer and tick (v) in the box.

Scales / Questions	1 Strongly agree	2 Agree	3 Slightly agree	4 Slightly disagree	5 Disagree	6 Strongly disagree
1. I **can answer** questions which I have memorized the answer in English class, e.g. 'Who inherited a fortune from Mr. Brown in this story?'						
2. I **can answer** questions which ask for information I have memorised, e.g. 'Recite the text.'						
3. I **can answer** questions where the answer can be found in the text, e.g. 'What did the old man say before he died in the text?'						
4. I **can answer** questions which ask for understanding of the materials learned in English class, e.g. 'What does this sentence mean?'						
5. I **can answer** questions which require a summary of an article or a story read or listened to in English class, e.g. "What is the main idea of this article?'.						
6. I **can answer** questions where I need to retell what I learned in English class, e.g. 'Describe what happened to World Trade Center in New York on September 11, 2001.'						

Scales / Questions	1 Strongly agree	2 Agree	3 Slightly agree	4 Slightly disagree	5 Disagree	6 Strongly disagree
7. I **can answer** questions where I need to apply things I have learned to a new context in English class, e.g. 'How is O zone layer related to global warming?'						
8. I **can answer** questions where I need to place myself in other situation, e.g. 'What would you do, if you were this person?'						
9. I **can answer** questions where I need to provide a new example in English class, e.g. 'Can you give me another example?'						
10. I **can answer** questions where I need to give reasons in English, e.g. 'Why do you make this decision**?**'						
11. I **can answer** questions which require comparison and contrast skills in English class, e.g. 'Can you tell the similarities and differences between these two characters in this story?'						
12. **I can respond to** questions where I need to analyse in English class, e.g. 'What are the strengths and weakness of merging these two companies?'						
13. In English class I **can respond to** questions where I need to think of a possible solution to a problem, e.g. 'How can you improve your English speaking ability?'						
14. I **can answer** questions which require creative skills, e.g. 'Can you make up a new ending for this English novel?'						
15. I **can respond to** questions where I need to design something new in English class, e.g. 'How can you design another new itinerary based on the budget we have?'						
16. **I can answer** questions which require assessment and reasoning skill, e.g. 'Which do you think is better and why?'						
17. I **can answer** questions where arguing ability is required and I need to elaborate my views and provide reasons or evidence in English class, e.g. 'What is your opinion towards this plan? Provide reasons or evidence.'						
18. I **can respond to** questions where I need to give my personal views, e.g. agree or disagree, and give reasons to support my views. e.g. 'Do you agree with this decision? Why?'						

Section 3

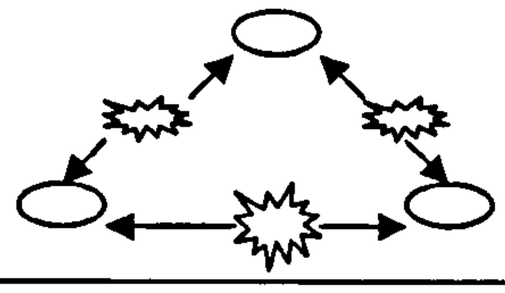
This section asks about **your belief about questioning-answering behaviour in English class**. There are six scales to choose from, 1= Strongly agree, 2= Agree, 3= Slightly agree, 4= Slightly disagree, 5= Disagree, 6= Strongly disagree. Please choose an appropriate answer and tick it in the box.

Questions \ Scales	1 Strongly agree	2 Agree	3 Slightly agree	4 Slightly disagree	5 Disagree	6 Strongly disagree
1. I can **talk freely in front of the class** when answering questions where the answers **can** be easily found in the text in English class.						
2. I can **express** my own opinions and ideas **freely in front of the class** when answering questions where the answers to the questions **can not** be easily found in the text in English class.						
3. In English class when the teacher asks a question where the answer **can not** be found in the text, the teacher usually gives us **enough time to think** about the question and answer it.						
4. After a **student expressed his/her opinions**, the teacher usually gives us **enough time to think** about what this student has said and allows us to comment or elaborate our own opinions in English class.						
5. After one student expressed his/her opinions, the teacher usually **asks** him/her **for additional information** in English class.						
6. 'Questions' which are **intellectually challenging** are interesting.						
7. I like to answer questions in English which are intellectually challenging.						
8. 'Questions' which are **intellectually challenging** provide me **more opportunities to talk in English**.						
9. 'Questions' which are intellectually challenging are useful in terms of **improving students' thinking** ability, e.g. analysis, creation, and evaluation.						

Section 4

This section asks about your opinions towards the intervention conducted. For question 1-7, there are six scales to choose from, 1= Strongly agree, 2= Agree, 3= Slightly agree, 4= Slightly disagree, 5= Disagree, 6= Strongly disagree. Please choose an appropriate answer and tick it in the box. For question 8-13, please choose an appropriate answer (s) and tick in the box.

The tasks we used in this intervention are as follows:

<table>
<tr><th>Task</th><th>Example</th></tr>
<tr><td>5Ws: a list of questions</td><td>Based on the 7 travelling packages provided, choose a suitable package for the role assigned to you. Provide reasons why this package is the most suitable one for the role and why other packages are not so suitable.</td></tr>
<tr><td>Odd One Out: Among the three, choose one which is different from the other two.</td><td></td></tr>
<tr><td>Make Up A Story: use 4 pictures to Make Up A Story</td><td><table><tr><td>Picture 1: Pont du guard bridge</td><td>Picture2: A naked man</td><td>Picture 3: The man swims in the river</td><td>Picture4: People laughed</td></tr></table></td></tr>
<tr><td>Guess What I Say:</td><td>A student expresses an idiom or proverb in English, others try to guess what the idiom or proverb is in Chinese.</td></tr>
</table>

Questions \ Scales	1 Strongly agree	2 Agree	3 Slightly agree	4 Slightly disagree	5 Disagree	6 Strongly disagree
1. Through practicing thinking skills tasks, I can talk more in English after this intervention.						
2. Through practicing thinking skills tasks, I can express my ideas more logically with reasons or evidence in English.						
3. Through practicing thinking skills tasks, I have more confidence in expressing my ideas in English, e.g. Even though I don't know certain grammar points or vocabulary I need, I still give it a try.						
4. Through practicing thinking skills tasks, my thinking skills have developed.						
5. I prefer to express my opinions within a small group rather than in front of the class.						
6. I like to listen to teacher's lecturing of vocabulary and grammar.						
7. I like to have more opportunities to practice English speaking.						

8. Which task is **intellectually challenging** to you and makes you think hard?
(You may tick more than one)
□ 5Ws □ Odd One Out □ Make Up A Story □ Guess What I Say □ None

9. Which task allows you to **talk more**? (You may tick more than one)
□ 5Ws □ Odd One Out □ Make Up A Story □ Guess What I Say □ None

10. Which task helps you **learn to express your ideas more logically**? (You may tick more than one)
□ 5Ws □ Odd One Out □ Make Up A Story □ Guess What I Say □ None

11. Which task helps you **gain more confidence** in expressing your ideas?
(You may tick more than one)
□ 5Ws □ Odd One Out □ Make Up A Story □ Guess What I Say □ None

12. Which task develops your **thinking skills**?
(You may tick more than one)
□ 5Ws □ Odd One Out □ Make Up A Story □ Guess What I Say □ None

13. Choose your **favourite** tasks. (You may tick more than one)
□ 5Ws □ Odd One Out □ Make Up A Story □ Guess What I Say □ None

後測問卷調查 **性別:**

本問卷旨在探討您在此研究後，對思考技巧、問與答的意見、及對此研究的看法，此為不記名問卷，並無對錯之分，您**最真實的回答**，就是最好的答案。

這份問卷包含四部份:
第一部份是關於在英文課堂中，您對所使用的思考技巧的看法及態度。
第二部份是關於在英文課堂中，您回答不同類型問題的能力。
第三部份是關於在英文課堂中，您對問與答的看法。
第四部份是關於您對此研究的看法。

第一部份

第一部份是關於在「**英文課堂**」**中，您對所使用的思考技巧的看法及態度。** 共六個選項可供選擇（1=非常同意，2=同意，3=少許同意，4=少許不同意，5=不同意，6=非常不同意），請選取一個合適的答案，並在格子裡打勾(v)。

問題 \ 程度	1 非常同意	2 同意	3 有點同意	4 有點反對	5 反對	6 非常反對
1 英文課中，用「**記憶力**」記住所學的東西，是很重要的。						
2 英文課中，使用「**背誦能力**」是很重要的。如，背誦英文課文。						
3 使用「**回想能力**」，回想起英文課中所學的東西是很重要的。						
4 英文課中，能用「**理解力**」理解文章或教材的意思，是很重要的。						
5 英文課中，能用「**做摘要的能力**」將所閱讀或聽到的英文文章簡短地說出其概要, 是很重要的。						
6 英文課中，能用「**敘述的能力**」將所閱讀或聽到的文章用英文重述一遍，是很重要的。						
7 英文課中，會使用「**應用所學的能力**」是很重要的。如，將所學到的閱讀技巧，應用在課後閱讀英文文章上。						
8 英文課中，能想像自己若是小說中的主角，會怎麼做，這種使用「**設身處地思考的能力**」，是很重要的。						
9 英文課中，當同學不了解我的意思時，我能用「**其他例子再解釋一次**」，這種能力是很重要的。						
10 英文課中，能「**說明原因及理由的能力**」是重要的。 例, 當我做一項決定時，我能說出做此決定的原因。						
11 英文課中，能使用「**比較的能力**」將兩件事做比較，並指出其異同點，是很重要的。 如，比較小說中兩角色間的差異。						
12 英文課中，能用「**分析的能力**」分析文章，是很重要的。如，分析文章中一項計劃的優缺點。						
13 英文課中，能使用「**解決問題的能力**」來解決問題，是很重要的。如，想出省錢的方法。						
14 英文課中，使用「**創造力**」是很重要的。如，重新改寫故事結局。						
15 英文課中，使用「**設計的能力**」是很重要的。如，設計兩週假期的旅行行程。						

問題 \ 程度	1 非常同意	2 同意	3 有點同意	4 有點反對	5 反對	6 非常反對
16 英文課中，會使用「**評估能力**」並提出理由及証據是很重要的。如，指出哪個旅行方案較適合學生，並提出理由及証據。						
17 英文課中，能使用「**辯論能力**」並提出理由及証據是很重要的。如，辯論校園中不該有服裝規定，闡述自己的意見，並針對其意見，提供相關的理由及証據。						
18 英文課中，會用「**判斷能力**」是很重要的。如，表達自己的立場，同意或不同意，並針對其立場，提供相關的理由及証據。						

第二部份

第二部份是關於在「**英文課堂**」中，**您回答不同類型問題的能力**，有六個選項可供選擇（1=非常同意，2=同意，3=少許同意，4=少許不同意，5=不同意，6=非常不同意）。請選取一個合適的答案，並在格子裡打勾(v)。

問題 \ 程度	1 非常同意	2 同意	3 有點同意	4 有點反對	5 反對	6 非常反對
1 英文課中，我**能用英文回答**有關我已記憶的東西。如，'故事中，誰從布朗先生那裡繼承了一大筆財富?'						
2 英文課中，我**能用英文回答**有關我已背誦的東西。如，'背出課文'。						
3 英文課中，我**能用英文回答**其答案是在課文裡的問題。如，'課文中，那老人死之前說了什麼?'						
4 英文課中，我**能用英文回答**需要解釋課文內容的問題。如，'這句子是什麼意思?'						
5 英文課中，我**能用英文回答**需做摘要的問題。如，'這篇文章的大意是什麼?'						
6 英文課中，我**能用英文回答**需敘述我所學的東西的問題。如，'描述 2001 年 9 月 11 日紐約世貿中心發生了什麼事。'						
7 英文課中，我**能用英文回答**需將所學運用到新的情境的問題。如，'臭氧層與全球暖化有什麼關係?'						
8 英文課中，我**能用英文回答**需假設自己在另一情境中的問題。如，'若你是這個人，你會怎麼做?'						
9 英文課中，我**能用英文回答**需提供新例子的問題。如，'你可以舉其他例子嗎?'						
10 英文課中，我**能用英文回答**需提供原因及理由的問題。如，'你為什麼要做這個決定?'						
11 英文課中，我**能用英文回答**需做比較的問題。如，'你可以比較故事中這兩個主角個性的異同點嗎?'						

問題 \ 程度	1 非常同意	2 同意	3 有點同意	4 有點反對	5 反對	6 非常反對
12 英文課中，我**能用英文回答**需做分析的問題。如，'這兩家公司合併，優缺點為何?'						
13 英文課中，我**能用英文回答**需思考解決方案的問題。如，'你如何改善英文口語能力?'						
14 英文課中，我**能用英文回答**需用創造力來思考的問題。如，'你能為這英文小說編寫一新結局嗎?'						
15 英文課中，我**能用英文回答**需用設計力來思考的問題。如，'依現有的旅行預算，你如何設計另一新的旅行路線?'						
16 英文課中，我**能用英文回答**需用評估能力的問題，辨別哪個較好，並提出理由及証據。如，'你認為哪個較好? 為什麼?'						
17 英文課中，我**能用英文回答**需用辯論能力的問題，闡述我的意見，並針對其意見，提供相關的理由及証據。如，'你對這計劃有何看法? 並提出理由及証據。'						
18 英文課中，我**能用英文回答**需用判斷能力的問題，表達我的立場，同意或不同意，並針對其立場，提供相關的理由及証據。如，'你同意這決定嗎? 為什麼?'						

第三部份

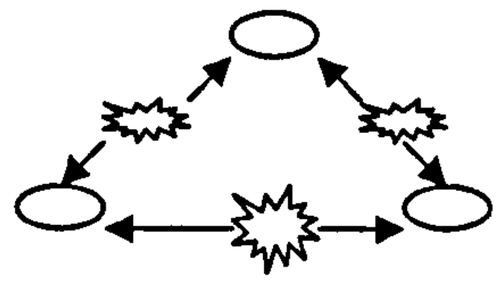
第三部份是關於在「**英文課堂**」**中，您對問與答的看法。** 有六個選項可供選擇（1=非常同意，2=同意，3=少許同意，4=少許不同意，5=不同意，6=非常不同意），請選取一個合適的答案，並在格子裡打勾(v)。

問題 \ 程度	1 非常同意	2 同意	3 有點同意	4 有點反對	5 反對	6 非常反對
1 英文課中，我可以**自在地在班上用英文回答**課本上**有答案**的問題。						
2 英文課中，我可以**自在地在班上用英文回答**課本上**沒有答案**的問題，這種答案是需自己想出來的。						
3 英文課中，當老師問的問題，其答案是**課本上找不到的**，老師通常給我們**足夠時間思考**，然後再回答。						
4 英文課中，**在學生表達其看法之後**，老師通常給我們**足夠時間思考**那位同學的說法，並讓我們去評論或表達其看法。						
5 英文課中，在學生表達其看法之後，老師通常會**詢問更進一步的相關訊息**。						
6 具**腦力挑戰**的問題，其答案是須靠自己動腦筋想出來的，這種問題是很有趣的。						
7 我**喜歡用英文回答**具腦力挑戰或思考的問題。						
8 具**腦力挑戰**的問題，提供我更多**用英文表達說話的機會**。						
9 具腦力挑戰的問題，在**改善學生思考技巧**很有幫助。如，改善運用、分析、創造、及評論等思考技巧。						

第四部份

第四部份是關於您**對此研究的看法**。 問題 1-7 有六個選項可供選擇（1=非常同意，2=同意，3=少許同意，4=少許不同意，5=不同意，6=非常不同意），請選取一個合適的答案，並在格子裡打勾(v)。問題 8-13 可複選，請選取合適的答案，並在格子裡打勾(v)。

此研究中我們所使用的口語練習，有下列四種

<table>
<tr><td>口語練習項目</td><td>題型</td></tr>
<tr><td>5Ws: 一些討論的問題</td><td>例如:
請選擇下列哪一個旅行方案較適合新婚夫婦， 並提供理由及證據。</td></tr>
<tr><td>Odd One Out: 不一樣的出局。
三個中，選出一個與其他兩個不一樣的</td><td></td></tr>
<tr><td>Make Up A Story: 用四張圖片編造一個故事</td><td><table><tr><td>圖片 1:
法國水道橋</td><td>圖片 2:
裸男</td><td>圖片 3:
河裡游泳</td><td>圖片 4:
人在笑</td></tr></table></td></tr>
<tr><td>Guess What I Say: 我說你猜</td><td>我說英文，你猜中文成語</td></tr>
</table>

問題 \ 程度	1 非常同意	2 同意	3 有點同意	4 有點反對	5 反對	6 非常反對
1. 藉由思考技巧的練習方式，我可以用英文說更多話。						
2. 藉由思考技巧的練習方式，我可以用英文將我的想法，表達的更有邏輯，並提出理由及証據。						
3. 藉由思考技巧的練習方式，我更有信心用英文表達我的想法。例如，雖然有些文法或單字不懂，我還是會試著說出來。						
4. 藉由思考技巧的練習方式，我的思考技巧，有被激發出來。						
5. 我較喜歡在小組討論中表達我的看法, 較不喜歡在全班前表達。						
6. 我喜歡老師多上些單字和文法。						
7. 我喜歡老師給我們更多用英文練習說話的機會。						

8. 下列哪幾項對你來說最具腦力挑戰，並需要努力思考? (可選一個或數個答案)
□ 討論的問題 □ 不一樣的出局 □ 編造故事 □ 我說你猜 □ 以上皆無

9. 下列哪幾項讓你更有機會講話? (可選一個或數個答案)
□ 討論的問題 □ 不一樣的出局 □ 編造故事 □ 我說你猜 □ 以上皆無

10. 下列哪幾項幫助你學習更有邏輯地表達你的想法? (可選一個或數個答案)
□ 討論的問題 □ 不一樣的出局 □ 編造故事 □ 我說你猜 □ 以上皆無

11. 在你表達意見時，下列哪幾項讓你獲得更多信心? (可選一個或數個答案)
□ 討論的問題 □ 不一樣的出局 □ 編造故事 □ 我說你猜 □ 以上皆無

12.下列哪幾項激發出你的思考技巧? (可選一個或數個答案)
□ 討論的問題 □ 不一樣的出局 □ 編造故事 □ 我說你猜 □ 以上皆無

13. 請選擇你喜歡的練習項目。(可選一個或數個答案)
□ 討論的問題 □ 不一樣的出局 □ 編造故事 □ 我說你猜 □ 以上皆無

Appendix 6 - An Example of Interview Transcription

T: 你喜歡這種教學法嗎？
(Do you like this teaching approach?)
L13: 我比較喜歡思考那方面
(I prefer thinking.)
T: 用英文回答思考的問題, 你可以嗎?
(Can you answer higher-order questions in English?)
L13: 我可以但我要想很久 因為我英文比較不好
(I can, but it took me a long time to think about it because my English is not good.)
T: 這次整個的教學法當中 你的思考有沒有比較頻繁?
(In this innovation, did you think more frequently?)
L13: 有阿
(yes, I did.)
T: 你會覺得這個對你的思考能力有沒有什麼引影響?
(Is there any impact on your thinking ability?)
L13: 有幫助 剛開始聽不懂在問什麼要想很久
(It helped. In the beginning of the innovation, I did not understand the question and it took me a long time to think.)
T: 那你會喜歡嗎？
(Do you like it then?)
L13: 可能要時間長一點才會開始喜歡.
(It might take a longer time for me to get to like it.)
T: 喔 那你覺得要一學期或兩學期嗎?
(Oh, then do you think it should last for one or two semesters?)
L13: 可能要從小培養.
(It might need to be taught since we were little.)
T: 說話次數有沒有增加?
(Did you talk more frequently?)
L13: 有
(Yes, I did.)
T:那你覺得這個對你口語來說有沒有什麼影響?
(Is there any impact on your speaking?)
L13: 就是要想一下再講 就會思考很多以前不會碰到那類型的單字 然後再把它講出來 然後才要去(unintelligible)之類的
(I needed to think for a while before I talked. I thought about words which I wouldn't use before. And then, I elaborated my ideas ...)
T: 你現在會用到之前...
(You now use words which ...)
L13: 都不會用到的字
(I wouldn't use before)
T: ok 你對這個會不會感到害怕 剛開始的時候
(Ok. Were you afraid of it in the beginning of the innovation?)
L13: 會阿
(Yes, I did.)
T: 有沒有遇到任何困難?
(Did you encounter any difficulty?)
L13: 就是要花時間想一下
(It took me some time to think.)
T: 謝謝你
(Thank you.)

Appendix 7 – Transcription System for Video Data

code	meaning
T	Teacher
L1; L2: etc.,	Identified learner
L1, L2	Two students at once
she is also cute no, I think	Utterance overlap between learners
=	Turn follows another without any pause
...	Pause of less than 4 seconds
(5)	Silence; length given in seconds
?	Rising intonation—question
(unintelligible)	Utterance can't be identified
(laughter)	Students laugh
Paul	Capitals are only used for proper nouns

Appendix 8 – Analysing Tools for Speaking

IELTS Speaking Band Descriptions (public version)

Band	Fluency and coherence	Lexical resource	Grammatical range and accuracy	Pronunciation
9	• speaks fluently with only rare repetition or selfcorrection; any hesitation is content-related rather than to find words or grammar • speaks coherently with fully appropriate cohesive features • develops topics fully and appropriately	• uses vocabulary with full flexibility and precision in all topics • uses idiomatic language naturally and accurately	• uses a full range of structures naturally and appropriately • produces consistently accurate structures apart from 'slips' characteristic of native speaker speech	• uses a full range of pronunciation features with precision and subtlety • sustains flexible use of features throughout • is effortless to understand
8	• speaks fluently with only occasional repetition or self-correction; hesitation is usually content-related and only rarely to search for language • develops topics coherently and appropriately	• uses a wide vocabulary resource readily and flexibly to convey precise meaning • uses less common and idiomatic vocabulary skilfully, with occasional inaccuracies • uses paraphrase effectively as required	• uses a wide range of structures flexibly • produces a majority of error-free sentences with only very occasional inappropriacies or basic/nonsystematic errors	• uses a wide range of pronunciation features • sustains flexible use of features, with only occasional lapses • is easy to understand throughout; L1 accent has minimal effect on intelligibility
7	• speaks at length without noticeable effort or loss of coherence • may demonstrate language-related hesitation at times, or some repetition and/or self-correction • uses a range of connectives and discourse markers with some flexibility	• uses vocabulary resource flexibly to discuss a variety of topics • uses some less common and idiomatic vocabulary and shows some awareness of style and collocation, with some inappropriate choices • uses paraphrase effectively	• uses a range of complex structures with some flexibility • frequently produces error-free sentences, though some grammatical mistakes persist	• shows all the positive features of Band 6 and some, but not all, of the positive features of Band 8
6	• is willing to speak at length, though may lose coherence at times due to occasional repetition, self-correction or hesitation • uses a range of connectives and discourse markers but not always appropriately	• has a wide enough vocabulary to discuss topics at length and make meaning clear in spite of inappropriacies • generally paraphrases successfully	• uses a mix of simple and complex structures, but with limited flexibility • may make frequent mistakes with complex structures, though these rarely cause comprehension problems	• uses a range of pronunciation features with mixed control • shows some effective use of features but this is not sustained • can generally be understood throughout, though mispronunciation of individual words or sounds reduces clarity at times
5	• usually maintains flow of speech but uses repetition, self-correction and/or slow speech to keep going • may over-use certain connectives and discourse markers • produces simple speech fluently, but more complex communication causes fluency problems	• manages to talk about familiar and unfamiliar topics but uses vocabulary with limited flexibility • attempts to use paraphrase but with mixed success	• produces basic sentence forms with reasonable accuracy • uses a limited range of more complex structures, but these usually contain errors and may cause some comprehension problems	• shows all the positive features of Band 4 and some, but not all, of the positive features of Band 6
4	• cannot respond without noticeable pauses and may speak slowly, with frequent repetition and self-correction • links basic sentences but with repetitious use of simple connectives and some breakdowns in coherence	• is able to talk about familiar topics but can only convey basic meaning on unfamiliar topics and makes frequent errors in word choice • rarely attempts paraphrase	• produces basic sentence forms and some correct simple sentences but subordinate structures are rare • errors are frequent and may lead to misunderstanding	• uses a limited range of pronunciation features • attempts to control features but lapses are frequent • mispronunciations are frequent and cause some difficulty for the listener
3	• speaks with long pauses • has limited ability to link simple sentences • gives only simple responses and is frequently unable to convey basic message	• uses simple vocabulary to convey personal information • has insufficient vocabulary for less familiar topics	• attempts basic sentence forms but with limited success, or relies on apparently memorised utterances • makes numerous errors except in memorised expressions	• shows some of the features of Band 2 and some, but not all, of the positive features of Band 4
2	• pauses lengthily before most words • little communication possible	• only produces isolated words or memorised utterances	• cannot produce basic sentence forms	• speech is often unintelligible
1	• no communication possible • no rateable language			
0	• does not attend			

Word Count System

1. Every word was tallied except fillers, repetition, reading questions on the handout, Chinese words and incomplete words. Examples are shown in the following table.
2. Contracted words, e.g. 'what's' was tallied as 2 words
3. Filler was defined in this study as sounds or words which were used to fill up conversational space and time.

	Example extracted from students' utterance	Word counted
1. Fillers (Er, uh, mmm, uh-huh, oh, ah, wow)	L8: mmm...Ken, first...	L8 got 2 words
	L11: weather L12: uh-huh	L11 got 1 word L12 got none
2. repetition (repetition here involves only self-repetitions)	L7: so maybe maybe one day... 'maybe' was repeated and was not counted.	L7 got 4 words
	Exception: L1: Oh, I think pretty woman is very important L3: no no no L3 said no for three times which was emphasising in disagreeing what L1 had said, not repeated. Therefore, three 'no's were counted under this circumstance.	L1 got 7 words L3 got 3 words
3.Reading questions on the handout	L2: ... choose three most important characters you think that a good friend should have....I think is funny, reliable and friendly. L2 was reading the question on the handout as shown in the underlined utterances, therefore, these words were not tallied.	L2 got 7 words
4. Chinese words	L5: But if you really love him...就真的 就真的愛 '就真的 就真的愛', these were Chinese words which were not counted.	L5 got 6 words
5.Incomplete words	L7: on the most important ...cri criteria.. 'cri' is incomplete, therefore it was not tallied.	L7 got 5 words

AS Units and Clauses Coding System
(Adopted from Foster et al. 2000)

Coding symbols

Boundary symbols	Meaning	Examples extracted from students' utterance
//	boundary of AS-units	//It's not easy to say.//
::	subordinate clause boundary within an AS-unit	//I think:: my ideal mate will be smart.//
{ }	Inside brackets are dysfluency , words like false starts, repetitions, and self-corrections	e.g1. false starts //{if I were} if he can understand me:: and have high EQ::, I will be with him very happy//
		e.g.2. repetitions // I think ::{my} my ideal mate have to responsible.//
		e.g.3. self-corrections //{I am a design} I will be a designer.//

An example of AS unit coding

Clause	Examples extracted from students' utterance	Number of AS units	Number of clauses
Independent clause : A clause including a finite verb	//You are out. //	1	1
Independent sub-clausal unit: Consisting either one or more phrases	e.g. 1: consisting ellipses elements L12: I don't think a couple or a lover should stay together all the time. L10: //not all the time//	1	1
	e.g.2 consisting irregular sentences //why?//	1	1
Subordinate clause: Consisting minimally of a finite or non-finite verb element plus at lease one other clause element.	e.g.1 //I think:: character is very important:: because I want someone:: who can caring me and:: understand me.//	1	5
	e.g. 2 //the most important is high EQ:: because I'm afraid of someone:: who have bad EQ// hobbies, I want someone:: who have same hobbies like me:: and we can do things together.//	2	6

Examples of topicalization, interruption, scaffolding, collaborated utterance and miscellaneous coding

	Examples extracted from students' utterance	Number of AS units	Number of Clauses
1. Topicalization: Topicalized noun phrases generally belong to the unit of which they are the topic	//Bad people he let me:: fear:: to approach him.//	1	3
2. Interruption: One speaker's utterance is interrupted by another speaker	L11: //if I want :: to do something successful::, I think:: L10: //you need friendship// L11: I need someone:; to support me.// L10 interrupted L11 by saying 'you need friendship'. However, L11 kept going on and finish the turn.	L11 got 1 AS unit and 5 clauses L10 got 1 AS unit and 1 clause	
3. Scaffolding: One speaker cannot access to the correct word and the word is provided by another speaker	L7: //but it will get {more}... L8:// better// L7: better and better ...and more..er 和平? L8, L9: //peace// L7: peace// L7 did not know how to say 'peace' in English, therefore, L8 and L9 provided the word in English.	L7 got 1 AS unit and 1 clause L8 got 2 units and 2 clauses L9 got 1 unit and 1 clause	
4. collaborated utterance One speaker starts the utterance , without finishing it, the other speaker takes the turn and complete the utterance.	L11: //so you don't// L10: //so you don't like him anymore//... L11 was trying to say something, before she finished the sentence L10 continued with L11's utterance and finished the sentence.	L11 was credited 1 AS unit and 1 clause. L12 was also credited 1AS unit and 1 clause.	
5. miscellaneous Incomplete utterance is not false start. It occurs because the speaker is still at planning stage and thinking. Before the sentence finishes, the turn is taken by the other speaker.	L10: //oh, that's ok//, but if he you can do it together:: maybe your huh...// L11: //I don't think :: a couple or a lover should er...stay together all the time//	L10 got 2 AS unit and 3 clauses	

	L11 did not finish the sentence and was still at planning stage. The turn, then, was taken by L10 before L11 finished the sentence.	

Utterance which does not take into account as a AS unit

	Examples extracted from students' utterance
Utterance spoken in Chinese	就是 輕鬆 自在 的生活
Reading questions listed on the handout	where the most ideal country you would like to live in based on the most important ... cri...criteria... criteria you choose provide with reasons of why these 6 criteria are the most important

Principles used to code AS units:
There are three levels of application of coding AS-units. Level one is to be used for a full analysis of all the data. Level two is to be used for highly interactional data and it excludes one-word minor utterances and echoic responses. Level three is for use in special cases where analysis of non-fragmentary AS-units is required (e.g. a sentence like 'yes, I hope so' won't be counted as a AS unit). This study adapted Level one which contained the analysis of full text. However, Chinese words spoken in this study was not taken into account.

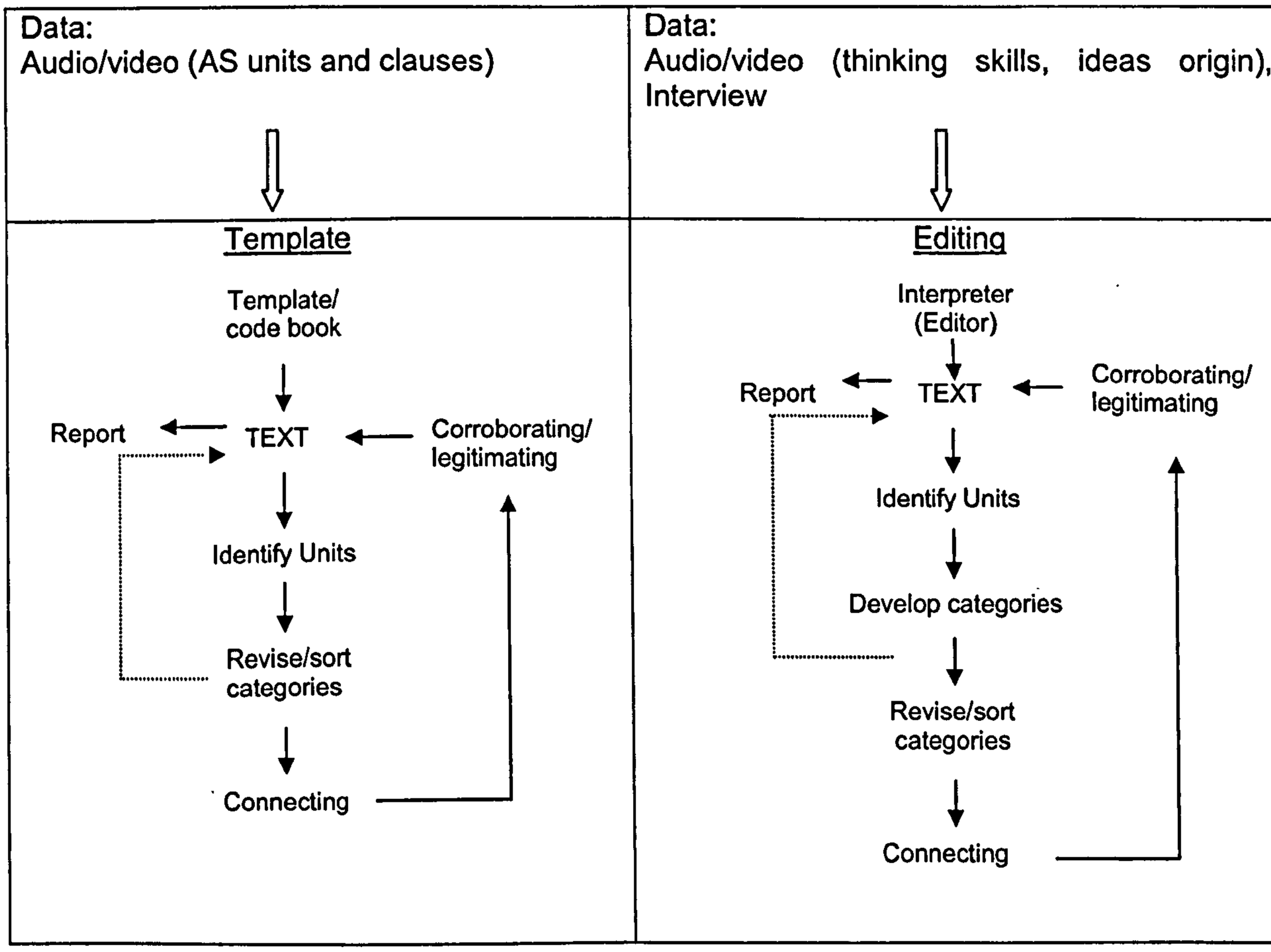
An example of AS Units and Clauses Coding

L12: //all the 3 question has {to}
L11: //to what?//
L12: to...answer and discuss//
L11: //{and} I think:: the three most important is...caring...{and...
L12: //caring//...
L11: caring,} supportive, and...reliable//
L12: //why you think :: supportive (unintelligible) {you think} is a good friend should have//...
L11: //mmm (7)...supportive// (4)
L10:// {I}, I can explain for you//...
L11: (laughter) //oh..ok//...
L10: // {because} I agree this...huh...{cha}...characters//...{we}...huh...we have to do...{maybe} something :: {we have to do...huh that something} we hadn't do it before,// {and...and} at that time, {we} have a friend :: {to support us, so...the friend}...huh... who is supportive :: is very important//...
L11: //{I think if}, if I want to:: do something {success}...successful::, {I}...I think::...huh...{some}=
L10: //{you}, you need friendship,//
L11: I need someone:: to support me//...I think:: it is important},//
L12: //do you think:: the caring character more like your husband or boyfriend//...{I think supportive...is er should ever}...I think :: supportive is... your boyfriend or girlfriend should have :: because {I think} =
L10: //but they can also use on friends//...
L12: {I think} friend maybe...er...comepass//...競爭怎麼講?
L10:// competition, //
L11: //competition//,
L12: //so I think a good friend//,
L10:// that is your enemy,//
L11: (laughter)
L12: //ok//,
L10: //but {is huh...we}...we can help each other,:: and... care each other,// so... the character caring is ok:: I think//...
L12: //and I think:: the most 2 important {charac}...of friend {is}...is... reliable and loyal...and//...
L11: (laughter) ..
L12: //I didn't choose//...
L10: //how do you tell... easy-going//...
L12: //easy-going...mmm...{ok}//=
L10: // {because you, you , you, you should, you have to...and} maybe you will stay together every day, every month, every weeks//...
L12: //{but, but if you}...mmm...for example, {if }I want :: to go {to eat} for lunch// and I ask you:: what do you want :: to eat//, and you mmm...whatever,// 然後 {and and}...I think:: I can't think//, I just {want to speak...what to eat, and I} need {your}...your answer,// {then} you just {all} say {whatever}, whatever, ::that make you annoyed,//
L11: // you think:: {if} ...if two people {are}... are easy-going::=
L12: //easy-going is//=
L11: how to make decision// ...
L12: // {no, no}, no...// easy-going is ok::, but not too over//...
//so...and you?//

Organizing Style of Analysis
(Adopted from Crabtree and Miller 1999)

Qualitative data collected in this study were analysed with content analysis. The coding methods used were Template and Editing (Crabtree and Miller, 1999). Template organizing style is that a template or code book is applied to the text being analysed with the intent of identifying the language units, such as words, phrases and utterance. If the text reveals inadequacies in the template, modifications are made and the text is re-examined. In this study template organizing style was used to examine the language units, AS units and clauses, in audio/video data .

On the other hand, editing organizing style is that the interpreter engages the text without a template and searches for meaningful units of text that both stand on their own and relate to the purpose of the study. The identified units are then sorted into categories. In this study, editing organizing style was used to analyse a) audio/video data: students' thinking skills and the function of ideas expressed, and b) interview data.



Appendix 9 – Analysing Tools for Thinking

Coding System of Six Thinking Levels

The coding system was based on Bloom's Taxonomy, cognitive domain (1956).

Categories	Examples extracted from students' utterance	Tally
1. Knowledge (rote recall of previously learned material, experience, or knowledge)	L2: Have you seen the snow view? L3: Yes, I have... L2 asked L3 whether she had seen snow view before. L3 replied by recalling of her experience and said 'yes, I have.'	One point was credited to L3 for Knowledge level.
	L6: Snow view Ia, snow view L5: Snow view L6 and L5 were both looking at the picture of snow view. They recognised and said that it was snow view.	L6 and L5 both got one point for Knowledge level.
2. Comprehension (the ability to make sense of the material, e.g. explaining what is meant or what is happening, summarizing, paraphrasing)	L10: If you tell your friend this secret, you want him or she to keep your secret... L12: If the friend is your good friend, he should be... L10: maybe you have many secrets, one of this secret you told your friend, and you don't want him to say it to other. L10 first stated that if one told a friend a secret, s/he wanted the friend to keep it as a secret. Then L10 further explained what she meant by saying 'maybe you have many secrets, one of this secret you told your friend, and you don't want him to say it to other.'	L10 got one point for Comprehension level.
	L12: If you are girl, what kind of boys you want to be your ideal mate L12 paraphrased the question in the handout. The original question was 'Suppose you are a female (girl/woman), think about what criteria are the most important for you when choosing an ideal mate.'	L12 got one point for Comprehension level.
	L1: may I choose 25? L2: 25?... L3:25? L1: Oh, I think pretty woman is very important. L3: no no no L2: it's 11 L1 thought item no. 25, pretty woman, was very important. However, L3 and L2 identify pretty woman was in item no.11.	Both L3 and L2 got one point for Comprehension level.

3. Application (Applied what was learned into a new context)	L7: ...I think the good friend should be, easy-going, just like Amy and Ken... L7 explained the character, easy-going, and gave examples for it.	On point was credited to L7 for Application level.
4. Analysis (the ability to break material into its component parts, e.g. comparing and contrasting, categorizing, hypothesizing, inferring)	L10: New Zealand have fresh airs, but Taiwan not... L10 was looking at the similarities and differences between Taiwan and New Zealand. The difference was that New Zealand had fresh air, but Taiwan did not.	L10 got one point for Analysis level.
	L9: If I can get a nice job, I can get more money. L9 hypothesized that if he can get a nice job, then he can get more money.	L9 got one point for Analysis level.
	L7: If you go to New York, maybe you'll get murdered one day, because there a lot of people can have gun. The context was that in New York lots of people had guns. So L7 inferred that if L8 went to New York, she might be murdered.	One point was credited to L7 for Analysis level.
5. Synthesis (the ability to put parts together, e.g. imaging, creating, problem solving)	L7: I like China, too... maybe one day I go there and I marry a husband, a rich guy. L7 imagined that one day she might go to China and marry to a Chinese guy.	L7 got one point for Synthesis level.
	L7: If you don't walking, you'll get fatter. L8: yeah, I know... But I can do less, not so much. L7 said to L8 if she does not walk, she will get fat. L8 came out with a solution to solve the problem of getting fat. She said she could do less, not so much which meant she would not take the public transportation too often.	L8 got one point for synthesis level.
6. Evaluation (the ability to judge the value of material based on specific criteria, e.g. evaluating, judging, assessing, appraising, criticizing)	L7: I want he be healthy... I will be a designer. The designer's health is not very good. He will take care of me... L7 considered health condition was an important criterion when choosing an ideal mate. She provided the reason by saying 'I will be a designer. The designer's health is not very good. He will take care of me.'	One point was credited to L7 for Evaluation level.
	L1: Pretty woman is very important... .... L3: Which country in the world is the most similar to the deal country you could create? ... L1: 'America'	L3 got one point for Evaluation level.

	... L1: I think Taiwan woman is pretty than American woman... L3: <u>But you think America is the ideal country, so...</u> L1 stated that the criterion for him to choose an ideal country to live in was somewhere with pretty woman. The ideal country he chose was America. However, L1 stated that 'Taiwan woman is pretty than American woman'. L3 then identified the inconsistency of L1's statement.	
	L12: I think appearance is important... ... L11: but I think appearance is not important. ... L11: appearance can change easy, the other character is hard to change L10: the appearance is not eternal... .. L10: maybe he have a car accident L12: so his face is...cry ... L10: so don't like him anymore... L12: yes, I understand...you mean ... L12: <u>so appearance is out</u> After the discussion of whether appearance was an important criterion, L12 made the decision that appearance should not be an important criteria by saying 'so appearance is out' based on the reasons discussed previously.	L12 got one point for Evaluation level.
7.Lower-order question (questions asked require the respondent to use lower order thinking, e.g. knowledge, comprehension, application)	L2: <u>have you seen the snow view?</u> L3: yes, I have... 'have you seen the snow view?' required L3 to recall his memory.	L2 got one point for the category of lower-order question
8. Higher-order question (questions asked require the respondent to use higher-order thinking, e.g. analysis, synthesis, evaluation)	L2: <u>do you want to go to New York?</u> L3: I think I want to go once 'do you want to go to New York?' required L3 to evaluate.	L2 got one point for the category of Higher-order question.

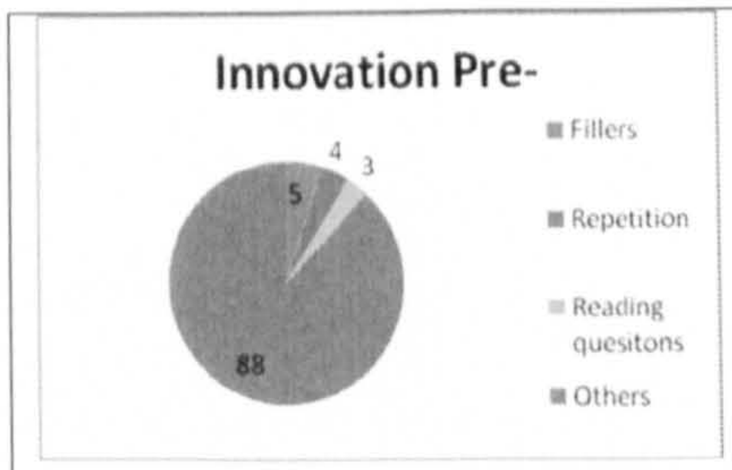
9. Any utterance coded as higher-order thinking, e.g. category 4, 5, and 6 appeared the second time was coded as lower-order thinking.

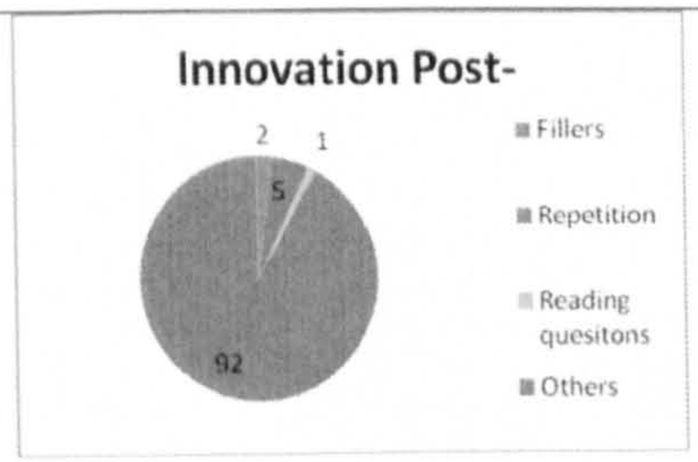
Coding System of the Function of HOTS

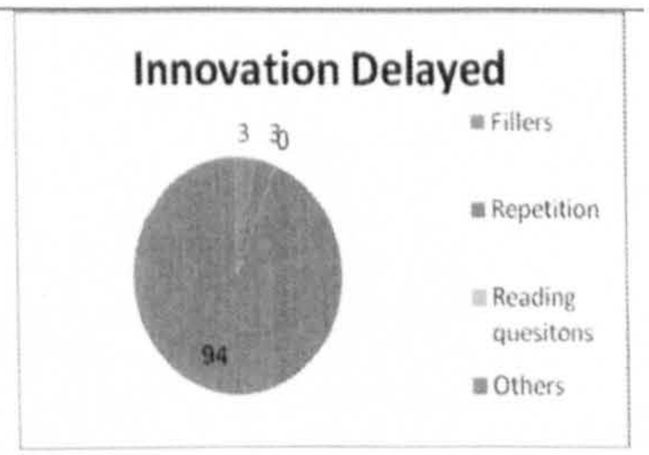
Categories	Examples extracted from students' utterance	Tally
A: HOTS ideas expressed based on answering the questions listed on the handout	L10: ok my turn, I choose character, hobbies, and soulmate...because I think if he can understand me, and have high EQ, I will be with him very happy L10 answered the question listed on the handout by saying 'I choose character, hobbies, and soulmate', and she further answering the question by explaining why she thought character was important. L10 expressed two ideas based on answering two different questions listed on the handout, therefore, two points were credited.	L10 got 2A
B: HOTS ideas expressed based on answering other's enquiry	L12: why you think supportive is a good friend should have ... L10: ...we have to do something we hadn't do it before, at that time have a friend who is supportive is very important... L12 asked the reason why the character of being supportive is important. L10 provided the reason, therefore, she was credited with one point.	L10 got 1B
C: HOTS ideas expressed based on directly build on or comment on other's opinion	L9: ...friendly is more like ...caring ... L7: but I think friendly is not the same like caring L7 commented on the idea generated by L9, therefore, she got one point.	L7 got 1C
	L10: maybe he have a car accident and his face is... L12: so his face is ...cry L12 collaborated with L10 to generate the idea. Under this circumstance, each of them got one point.	L10 and L12 both got 1C

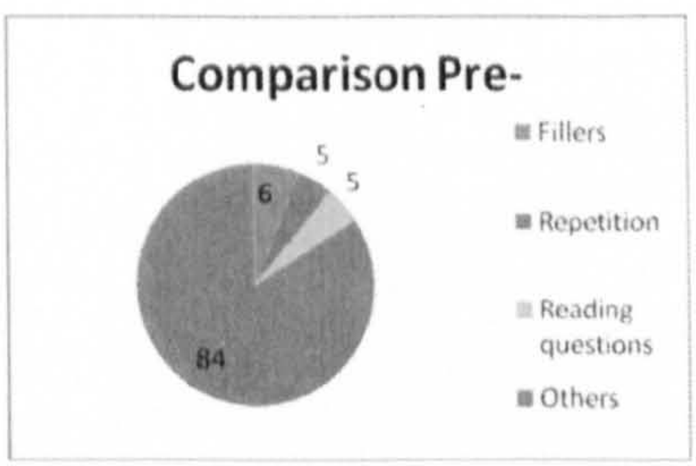
Appendix 10 – Analysis of the Amount of Words

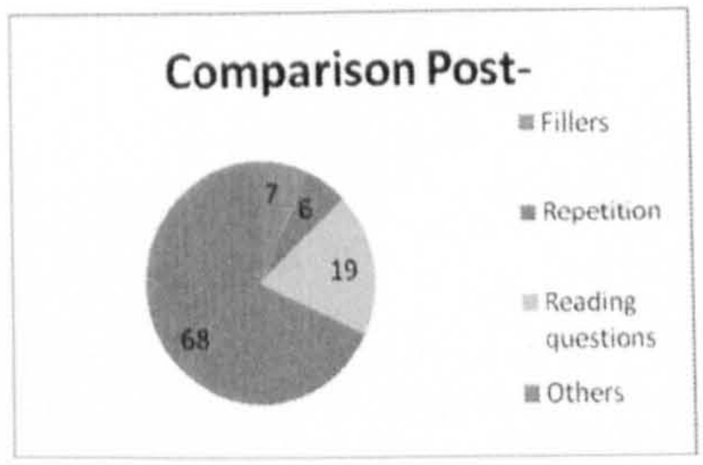
An Analysis of Percentage of Fillers, Repetition and Reading Questions

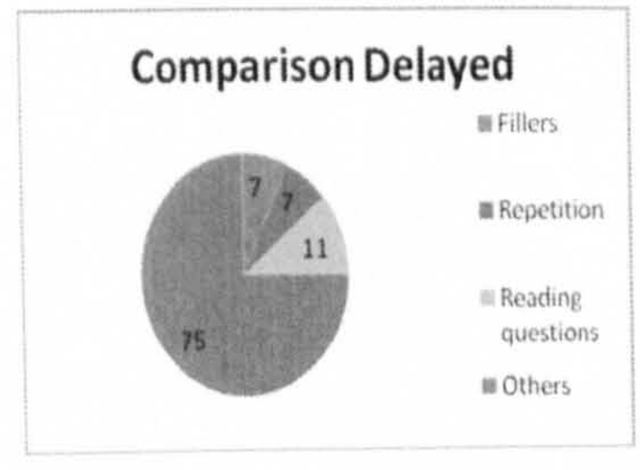




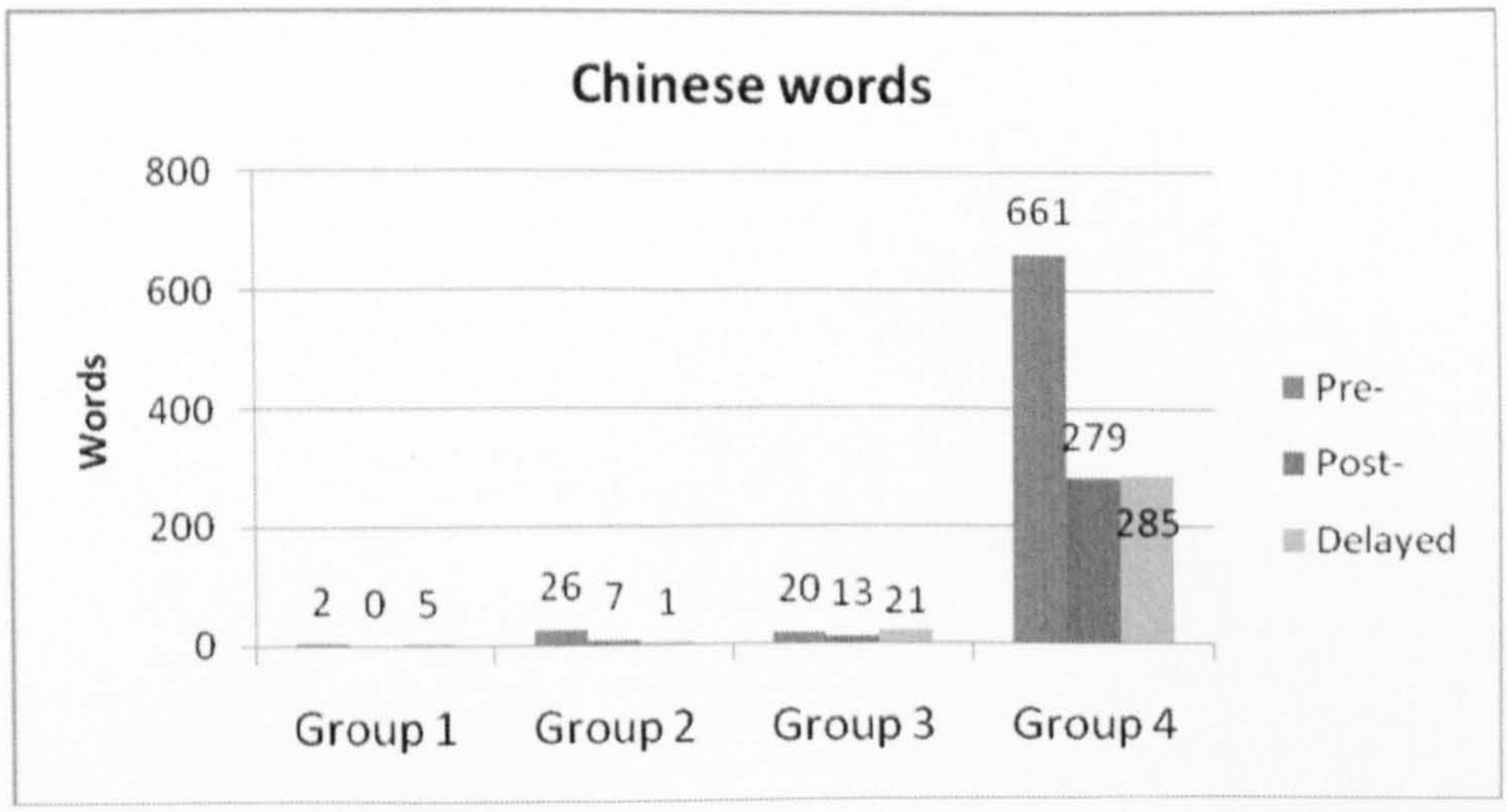








An analysis of the Total Number of Chinese Words



Group 1 and 2 were in the innovation class. Group 3 and 4 were in the comparison class.