A COMPARATIVE STUDY BETWEEN THE C-TEST AND THE NC-TEST AND BEWEEN THE MC-TEST AND THE NMC-TEST, USING IDENTICAL TEXTS

MONTARAT RUNGRUANGTHUM

A THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR
THE DEGREE OF MASTER OF ARTS
(APPLIED LINGUISTICS)
FACULTY OF GRADUATE STUDIES
MAHIDOL UNIVERSITY
2005

ISBN 974-04-6197-2
COPYRIGHT OF MAHIDOL UNIVERSITY

Copyright by Mahidol University

Thesis
Entitled

A COMPARATIVE STUDY BETWEEN THE C-TEST AND THE NC-TEST AND BETWEEN THE MC-TEST AND THE NMC-TEST, USING IDENTICAL TEXTS

Montarat Rungruangth
Miss. Montarat Rungruangthum
Candidate

Patama Kittidhaworn
Lect. Patama Kittidhaworn,
Ed.D.
Major-Advisor

Willi M. Martin
Mr. William M. Martin,
M.A.
Co-Advisor

Ram Hun
Assoc. Prof. Rassmidara Hoonsawat,
Ph.D.
Dean
Faculty of Graduate Studies

W. Puntai
Asst. Prof. Wiwat Puntai, Ph.D.
Chairman
Master of Arts Programme in
Applied Linguistics
Faculty of Science

Copyright by Mahidol University

Thesis
Entitled

A COMPARATIVE STUDY BETWEEN THE C-TEST AND THE NC-TEST AND BETWEEN THE MC-TEST AND THE NMC-TEST, USING IDENTICAL TEXTS

was submitted to the Faculty of Graduate Studies, Mahidol University
for the degree of Master of Arts (Applied Linguistics)
on
May 27, 2005

Miss Montarat Rungruangthum
Candidate

Lect. Patama Kittidhaworn,
Ed.D.
Chair

Mr. William M. Martin,
M.A.
Member

Lect. Somsak Boonsathorn,
Ph.D.
Member

Assoc. Prof. Rassmidara Hoonsawat,
Ph.D.
Dean
Faculty of Graduate Studies
Mahidol University

Prof. Amaret Bhumiratana,
Ph.D.
Dean
Faculty of Science
Mahidol University

Copyright by Mahidol University

ACKNOWLEDGEMENTS

I would like to say many thanks to Dr. Patama Kittidhaworn, my thesis advisor for her great kindness, guidance, encouragement, and valuable recommendations about this study. Also, I would like to thank Lect. Richard R. Marshal and Lect. William M. Martin, my co-advisors for editing and proofreading my thesis research. Without their assistance, this research would not have been achieved.

I am also grateful to Dr. Somsak Boonsathorn, Associate Dean for Academic Affairs in the School of Language and Communication at the National Institute of Development Administration (NIDA) for his kindness, valuable advice, assistance, and for being my external examiner. His research article hightens my interest in language testing.

I am indebted to all the lecturers and all staff members of the Foreign Language Department for their assistance in facilitating data collection. My thanks also go to all the first-year undergraduate science students for participating in this study.

Above all else, I would like to express my deepest gratitude to my parents, my sister, my brother, and my best friends. Thank you for always standing by me and never stopping believing so that I could complete this test of endurance.

Montarat Rungruangthum

Copyright by Mahidol University

A COMPARATIVE STUDY BETWEEN THE C-TEST AND THE NC-TEST AND BETWEEN THE MC-TEST AND THE NMC-TEST, USING IDENTICAL TEXTS.

MONTARAT RUNGRUANGTHUM 4436732 SCAL/M

M.A. (APPLIED LINGUISTICS)

THESIS ADVISORS: PATAMA KITTIDHAWORN, Ed.D., WILLIAM MARTIN, M.A.

ABSTRACT

This study aims to investigate the four types of cloze tests by comparing the mean scores on the C-Test with those on the NC-Test, and the mean scores on the MC-Test with those on the NMC-Test in measuring the English language proficiency of 110 first-year undergraduate science student studying at the Faculty of Science, Mahidol University. These target students with high-language ability (55) and low-language ability (55) were divided into four sub-groups. Each sub-group was required to take one type of the four cloze tests at the start of their English classes. The test-taking strategies, based on Sasaki's framework (2000), were also studied to see how these target students responded to each test type. The statistical analyses used in the study were the mean, the Independent Sample t-test, Pearson's Product Moment Correlation, and the Kuder-Richardson 21 Formula (K-R 21).

The results indicate that there were no statistically significant differences at the .05 level of significance ($p <$. 05) between the C-Test and the NC-Test in both high- and low-language-ability groups. There is a statistically significant difference at $p < .05$ between the MC-Test and the NMC-Test only in the high-language-ability group. So the NC-Test may be used as a substitute for the C-Test in measuring the English language proficiency of EFL tertiary students. The MC-Test and the NMC-Test are most acceptable to use for measuring the English language proficiency of high-language-ability students, while the NMC-Test seems to be less suitable for measuring the English language proficiency of low-language-ability students than the original MC-Test. The test-taking strategies most frequently used by volunteer students from each sub-group were *Within Clause; Across Clause, Within Sentence; Across Sentences, Within Paragraph;* and *Guessing*.

The findings suggest that language teachers who intend to use these four cloze tests should carefully consider the deletion techniques and the selection of the text content suitable for the students' language ability levels.

KEY WORDS: ENGLISH LANGUAGE TESTING/ THE C-TEST/ THE NC-TEST/ THE MC-TEST/ THE NMC-TEST/ TEST-TAKING STRATEGIES

P. 143 ISBN 974-04-6197-2

Copyright by Mahidol University

การศึกษาเชิงเปรียบเทียบระหว่างแบบทดสอบC-TEST กับ NC-TEST และ แบบทดสอบMC-TEST กับ NMC-TESTโดยใช้เนื้อความเดียวกัน (A COMPARATIVE STUDY BETWEEN THE C-TEST AND THE NC-TEST AND BETWEEN THE MC-TEST AND THE NMC-TEST, USING IDENTICAL TEXTS)

มณฑรัตน์ รุ่งเรืองธรรม 4436732 SCAL /M
ศศ.ม. (ภาษาศาสตร์ประยุกต์)
คณะกรรมการควบคุมวิทยานิพนธ์ : ปัทมา กิตตถาวร, Ed.D., WILLIAM M. MARTIN, M.A.

บทคัดย่อ

การศึกษานี้มีวัตถุประสงค์เพื่อเปรียบเทียบการใช้แบบทดสอบโคลซสี่ชนิดโดยใช้คะแนนเฉลี่ยที่ได้จากแบบทดสอบ C-Test กับ NC-Test และที่ได้จากแบบทดสอบ MC-Test กับNMC-Test เพื่อวัดความสามารถในการใช้ภาษาอังกฤษของนักศึกษาสาขาวิทยาศาสตร์ ระดับปริญญาตรีชั้นปีที่หนึ่ง คณะวิทยาศาสตร์ มหาวิทยาลัยมหิดล จำนวน 110 คน โดยแบ่งเป็นสองกลุ่มใหญ่ๆ คือกลุ่มที่เก่งภาษาอังกฤษ (55 คน) กับ กลุ่มที่อ่อนภาษาอังกฤษ (55 คน) และแต่ละกลุ่มยังแบ่งออกเป็นสี่กลุ่มย่อย เพื่อทำแบบทดสอบแต่ละชนิดเพียงแบบเดียว นอกจากนี้ยังได้ศึกษากลยุทธในการทำแบบทดสอบโคลซแต่ละชนิด โดยใช้เกณฑ์ของ Sasaki (2000) การวิเคราะห์ข้อมูลใช้ค่าเฉลี่ย การทดสอบความแตกต่างของค่าเฉลี่ยของตัวแปรที่เป็นอิสระต่อกัน สัมประสิทธิ์สหสัมพันธ์ของเพียร์สัน และค่าความเชื่อถือได้โดยใช้สูตร Kuder-Richardson 21 (K-R 21)

ผลการศึกษาพบว่าไม่มีความแตกต่างอย่างมีนัยความสำคัญทางสถิติที่ระดับ .05 ระหว่างผลคะแนนที่ได้จากแบบทดสอบ C-Test กับ NC-Test ทั้งกลุ่มที่เก่งและอ่อนภาษาอังกฤษ และที่ได้จากแบบทดสอบ MC-Test กับ NMC-Test เฉพาะกลุ่มที่เก่งภาษาอังกฤษเท่านั้น ดังนั้นจึงอาจสรุปได้ว่าแบบทดสอบ NC-Test ใช้วัดความสามารถในการใช้ภาษาอังกฤษของนักศึกษากลุ่มนี้ได้เช่นเดียวกับแบบทดสอบ C-Test ส่วนแบบทดสอบ MC-Test ใช้วัดความสามารถในการใช้ภาษาอังกฤษของกลุ่มที่อ่อนภาษาอังกฤษได้เหมาะสมกว่าแบบทดสอบ NMC-Test. จากผลการสัมภาษณ์อาสาสมัครทั้งกลุ่มที่เก่งและอ่อนภาษาอังกฤษ พบว่ากลยุทธที่ใช้บ่อยที่สุดได้แก่ *Within Clause; Across Clause, Within Sentence; Across Sentences, Within Paragraph* และ *Guessing* ดังนั้นผู้สอนภาษาอังกฤษสามารถนำแบบทดสอบโคลซทั้งสี่ชนิดไปใช้วัดความสามารถในการใช้ภาษาอังกฤษของนักศึกษากลุ่มดังกล่าวได้ โดยพิจารณาเลือกใช้เทคนิคการเว้นคำและชนิดของคำให้เหมาะสมกับระดับความสามารถในการใช้ภาษาอังกฤษของนักศึกษา

143 หน้า ISBN 974-04-6197-2

Copyright by Mahidol University

CONTENTS

		Page
ACKNOWLEDGEMENT		**iii**
ABSTRACT		**iv**
LIST OF TABLES		**viii**
CHAPTER		
I	**INTRODUCTION**	**1**
	1.1 Background and rationale of the study	1
	1.2 Purpose of the study	8
	1.3 Significance of the study	8
	1.4 Scope and limitations of the study	9
	1.5 Definitions of terms	9
II	**LITERATURE REVIEW AND RELATED RESEARCH**	**11**
	2.1 Background of the cloze test	11
	2.2 The C-Test vs. the NC-Test	18
	2.3 The MC-Test vs. the NMC-Test	21
	2.4 Test-taking strategies	24
III	**METHODOLOGY**	**28**
	3.1 Subjects	28
	3.2 Research instruments	29
	3.2.1 The construction of the C-Test	34
	3.2.2 The construction of the New C-Test	35
	3.2.3 The construction of the Modified C-Test	36
	3.2.4 The construction of the New Modified C-Test	37
	3.3 Scoring procedure	38
	3.4 Piloting four different forms of cloze Tests	38
	3.5 Interview	40
	3.6 Data collection	42

Copyright by Mahidol University

CONTENTS (CONTS.)

3.7 Data Analysis 43

IV RESULTS 45

4.1 Results of the original MC-Test and the NMC-Test 45

4.2 Results of the original C-Test and the NC-Test 46

4.3 Results of using third-word deletion 48

4.4 Results of cloze test-taking strategies 54

V DISCUSSION 61

5.1 Comparison of the MC-Test and the NMC-Test 61

5.2 Comparison of the C-Test and the NC-Test 63

5.3 Effects of using third-word deletion 65

5.4 Cloze test-taking strategies used by the volunteer students 67

VI CONCLUSION 71

6.1 Conclusion 71

6.2 Recommendations for further studies 73

REFERENCES 75

APPENDIX 82

BIOGRAPHY 143

Copyright by Mahidol University

LIST OF TABLES

Table		Page
Table 1	Discrete-point approach and integrative approach	3
Table 2	Numbers of the high and low groups taking each type of the cloze test	29
Table 3	The average number of sentences and syllables of the three selected reading passages shown in Example 15 and the plotted place on Fry's readability	32
Table 4	Sources and estimated readability levels of the selected passages used in the present study	33
Table 5	The reliability of the four new forms of the cloze tests, using the K-R 21	40
Table 6	Cloze test-taking strategies categorization by Sasaki (2000, p.95)	42
Table 7	Comparison of the scores on the original MC-Test and the NMC-Test within the high-language-ability and low-language-ability groups	46
Table 8	Comparison of the scores original C-Test and the NC-Test within the high-language-ability and low-language-ability groups	47
Table 9	Comparison of the mean scores of the four cloze tests between the High-language-ability and low-language-ability groups	48
Table 10	Item discrimination of the four types of cloze tests	49
Table 11	The C-Test items with low discrimination power	50
Table 12	The NC-Test items with low discrimination power	51
Table 13	The MC-Test items with low discrimination power	52
Table 14	The NMC-Test items with low discrimination power	53
Table 15	The cloze test-taking strategies used by the high-language-ability and low-language ability groups	55
Table 16	The correlation of ratings cloze test-taking strategies in the high-language-ability and low-language-ability groups indicated by two inter-raters	56

Copyright by Mahidol University

LIST OF TABLES (CONTS.)

Table		Page
Table 17	Test-taking strategies for the C-Test transcribed as recorded	57
Table 18	Test-taking strategies for the NC-Test transcribed as recorded	58
Table 19	Test-taking strategies for the MC-Test transcribed as recorded	59
Table 20	Test-taking strategies for the NMC-Test transcribed as recorded	60
Table 21	Number of content words and functional words in the four types of cloze test	66

Copyright by Mahidol University

CHAPTER I
INTRODUCTION

1.1 Background and Rationale of the Study

Language assessment is an instrument for language teachers to identify the students' strengths and weaknesses in language learning, to place the student into a program and to measure the use of English in four basic skills (reading, writing, listening, and speaking). The assessment can be done by such methods as tests, interviews, or observations. For language teachers, the tests provide "evidence of the results of learning and instruction, and hence feedback on the effectiveness of the teaching program" (Bachman & Palmer, 1996, p. 8). The test results enable the students to develop their performance in language learning effectively. In addition, it is very important to select the most suitable language tests which respond to the specific goals of teaching. Language teachers should also understand the functions and the characteristics of the language tests thoroughly. Many practitioners and researchers in language testing (Bachman & Palmer, 1996; Brown, 1996; Hughes, 2003; McNamara, 2000) categorize four kinds of language tests based on the test purposes and functions as follows:

(1) *Proficiency Tests* are designed to measure general language skills, including speaking, listening, reading and writing. In addition, proficiency tests generally help teachers to "set up entrance and exit standards for a curriculum" (Brown, 1996, p. 9). For instance, the Test of English as a Foreign Language (TOEFL) and International English Language Testing System (IELTS) are currently used by many universities where English language proficiency is required.

(2) *Achievement Tests* are aimed at the degree of learning or how much progress the students have made (McNamara, 2000). So achievement tests are directly relevant to the goals of learning and instruction. These tests can be given in the middle or at the end of the program (Hughes, 2003; McNamara, 2000).

(3) *Diagnostic Tests* are established to analyze the students' strengths and weaknesses in the learning process (Brown, 1996; Hughes, 2003). These tests are conducted at the beginning of the program (Brown, 1996).

Copyright by Mahidol University

(4) *Placement Tests* are focused on screening the students to see whether they can study in a program and grouping the students in the same level of language proficiency (Hughes, 2003). Hence, the results of these tests will enable the teachers to accurately place the students entering any institution or program (Bachman & Palmer, 1996).

In addition to a clear understanding of the functions and the characteristics of language tests, language teachers have to understand the construction of those tests. There are two approaches which have an influence on test construction: the discrete-point approach and the integrative approach (Hughes, 2003). For the discrete-point approach, language teachers view each language component separately, measuring one language skill at a time, such as testing grammar or vocabulary (Brown, 1996; McNamara, 2000). In language testing, discrete-point tests emphasize language form rather than language use (McNamara, 2000). However, the discrete-point test results focusing on a single language component are inadequate to determine the student's language proficiency (Jitendra & Rohena-Diaz, 1996). As a consequence, Oller (1979) suggests that teachers should construct language tests using the integrative approach instead.

In the integrative approach, the language teachers view language as a whole, emphasizing both productive and receptive skills (Brown, 1996; Hughes, 2003; McNamara, 2000). Integrative tests, such as cloze, dictation, writing an essay, and interview, can measure several skills simultaneously (Brown, 1996; Hughes, 2003). Moreover, integrative tests are suitable for assessing language proficiency and communicative skills (Brown, 1996; McNamara, 2000). McNamara (2000) contends that integrative tests take a lot of time to construct and score, as shown in Table 1. However, cloze tests are reported to be less time consuming, easier to score, and more reliable in measuring students' English language proficiency (Oller, 1979).

Copyright by Mahidol University

Table 1: Discrete-point approach and integrative approach

(Based on Jitendra & Rohena-Diaz, 1996; McNamara, 2000)

Approach	Focus	Advantages	Disadvantages
Discrete-point	a single part of language can be tested separately.	- economical and quick to score - suitable for assessing certain language capabilities	- overemphasizes correctness and structures - cannot measure communicative language skills
Integrative	measures various skills at the same time and views language as a whole.	- suitable to assess both productive and receptive skills	- time-consuming and tends to be expensive - difficult to score

The cloze test was initiated by Taylor (1953, cited in Oller & Conrad, 1971). Originally, there were two kinds of cloze tests: a *rational cloze* and a *random cloze* (see Example 1). The former refers to the deletion of specific types of words in a selected passage, such as prepositions or articles. The latter deals with a consistent deletion of every n^{th} word, such as every fifth or seventh word. The student's task is to fill in the deleted part in the cloze passage. Cloze tests can measure grammatical structure, written expression and vocabulary as well as reading comprehension (Steinman, 2002). In addition, some studies (Aitken, 1977; Oller & Conrad, 1971; Oller, 1979; Stubbs & Tucker, 1974) indicate that the cloze test is an effective instrument which is reliable and valid to measure English language proficiency. But the different deletion rates have an effect on the validity and the measurement of the cloze test (Alderson, 1979, 1980, 1983, 2000). Klein-Braley (1997) adds that the deletion rates used in cloze tests require long passages. If a cloze test with the deletion of every fifth word provides 50 items, the text length should be at least 250 words (Oller, 1979). This problem has led to the development a new form of the cloze test which is called the C-Test.

Copyright by Mahidol University

Example 1: Rational and random cloze procedures

RATIONAL CLOZE

Choose *a*/*an*/*the* or no article at all. Then put */a/, /an/, or /the/* in each blank where necessary.

People today are quite astonished by **1)**_____ rapid improvements in medicine. Doctors are becoming more specialized, and new drugs are appearing on **2)**_____ market daily. At **3)**_____ same time, **4)**_____ people are dismayed by **5)**_____ inaccessibility of doctors when they are needed. Whereas doctors' fees are constantly on **6)**_____ rise, **7)**_____ quality of medical care has reached **8)**_____ abysmal low.

(Adapted from Cohen, 1994, p. 234)

RANDOM CLOZE

Fill in the missing words

People today are quite astonished by the rapid improvements in medicine. Doctors **1)**_____ becoming more specialized, and **2)**_____ drugs are appearing on the **3)**_____ daily. At the same time, **4)**_____ are dismayed by the inaccessibility **5)**_____ doctors when they are needed. **6)**_____ doctors' fees are constantly on **7)**_____ rise, the quality of medical **8)**_____ has reached an abysmal low.

(Adapted from Cohen, 1994, p. 234)

"The C-Test", one of the new cloze tests, was constructed by Raatz and Klein-Braley (1981) in order to see if it could be more effective than the original cloze tests in measuring the students' English language proficiency. The construction of the C-Test is based on the same principle as that of the cloze test; however, only the *second half* of every *second word* is deleted as can be seen in Example 2. In the C-Test, if the deleted word contains an even number of letters, the second half of this word will be deleted, such as "**exper i e n c e**" (10 letters). For a word with an odd number of letters, its larger part must be deleted, such as "**th e r e**" (5 letters). Moreover, many research studies indicate that the C-Test is more effective and more reliable than the original cloze (Connelly, 1997; Dörnyei & Katona, 1992; Klein-Braley, 1985, 1997), and yet, Dörnyei and Katona (1992) report that the C-Test is too difficult for non-native students studying a target language such as English.

Copyright by Mahidol University

As a result, Thongsa-nga (1998) adopted the original C-Test to make it suitable for Thai students studying English as a foreign language. Imitating the C-Test construction, Thongsa-nga (1998) proposed "the New C-Test (the NC-Test)" by deleting the *second half* of every *third word* in order to provide more clues for the non-native test takers, as can be seen in Example 2. According to the investigation of Thongsa-nga (1998), the NC-Test is employed as a proficiency test for non-native students at a secondary school level. The findings reveal that the NC-Test is reliable to assess the English language proficiency of these Thai Mathayomsuksa Six students. As far as this researcher has been able to establish, there has been no research investigating the use of the NC-Test for non-native university students in Thailand. So the present study is designed to examine the similarities and the differences in using the C-Test and the NC-Test in measuring the English language proficiency of first-year Thai undergraduate students.

Example 2: The C-Test and the NC-Test

THE C-TEST

Many foreigners find that Thailand is a very pleasant place to have a holiday. They **disc** _ _ _ _ that **th** _ _ _ are **ma** _ _ interesting **thi** _ _ _ to **d** _ and **t** _ see. **Th** _ _ say **th** _ _ the **bea** _ _ _ _ are **cl** _ _ _ and **t** _ _ scenery **i** _ beautiful. **Ma** _ _ say **th** _ _ the **hot** _ _ _ are **exce** _ _ _ _ _ and **n** _ _ too **expe** _ _ _ _ _. They **exper** _ _ _ _ _ with **diff** _ _ _ _ _ kinds **o** _ Thai **fo** _ _ that **i** _ tastes **deli** _ _ _ _ _.

(Boonsathorn, 1990, p. 48)

THE NC-TEST

There is a dark shadow over schools and colleges where students are now facing the enormous problem of drugs. There seems **t** _ be an **incr** _ _ _ _ in the **u** _ _ of alcohol, **tob** _ _ _ _ and other **dr** _ _ _ by students. **Sch** _ _ _ _ and colleges **ha** _ _ to try **t** _ deal with _ growing number **o**_ students victimized, **eit** _ _ _ directly or **indir** _ _ _ _ _ by drug **u** _ _ or drug **dea** _ _ _ _.

(Thongsa-nga, 1998, p. 43)

Copyright by Mahidol University

Another form of the cloze test, **"**the Modified C-Test (the MC-Test)**"**, also known as "the X-Test", was invented by Boonsathorn in 1987 (cited in Boonsathorn, 1990, p. 46). For the MC-Test, the *first half* of every *second word* is deleted (see Example 3). In the MC-Test, if the total number letters of the deleted word is an even number, the first half of this word will be deleted, such as **"d i s a gree**" (8 letters). For a word with an odd number of letters, its larger part will be deleted, such as "**o t h er**". According to Boonsathorn (1987), the first half deletion in the MC-Test compares with the C-Test. His study reports that the MC-Test is more difficult and discriminates better than the C-Test. Some research findings show that the MC-Test has high reliability and validity and can be used with advanced students (Köberl & Sigott, 1996; Prapphal, 1994; Sigott & Köberl, 1993; Wonghiransombat, 1998). So the MC-Test should be further investigated to see its strengths and weaknesses in assessing English language skills. The MC-Test can be an alternative type for a better assessment of the English language proficiency of Thai undergraduate students, although the study of Sigott and Köberl (1993) claims that the MC-Test is more difficult for non-native speakers.

Wonghiransombat (1998) then proposed "the New Modified C-Test (the NMC-Test)" in order to make the original MC-Test appropriate for non-native students (p. 23). The construction of the NMC-Test is based on the same principle as the MC-Test; however, the *first half* of every *third word* is deleted to provide more clues as shown in Example 3. In addition, Wonghiransombat (1998) reports that the NMC-Test with the *third* starting point, or the third-word deletion, is easier and has better discrimination than the original MC-Test. Her study, the only research done in Thailand to examine the use of the MC-Test and the NMC-Test at the postgraduate level, also shows that the NMC-Test can be utilized to measure English language proficiency of Thai postgraduate students. Therefore, the present study is also aimed at examining the similarities and the differences in using the original MC-Test and the NMC-Test in measuring English language proficiency of the Thai undergraduate students.

Copyright by Mahidol University

Example 3: The MC-Test and the NMC-Test

THE MC-TEST

Many foreigners find that Thailand is a very pleasant place to have a holiday. They _ _ _ _ **over** that _ _ _ **re** are _ _ **ny** interesting _ _ _ **ngs** to _ **o** and _ **o** see. _ _ **ey** say _ _ **at** the _ _ _ _ **hes** are _ _ _ **an** and _ _ **e** scenery _ **s** beautiful. _ _ **ny** say _ _ **at** the _ _ _ **els** are _ _ _ _ _ **lent** and _ _ **t** too _ _ _ _ _ **sive**.

(Boonsathorn, 1990, p. 49)

THE NMC-TEST

One of the most vital aspects of an orientation program is training supervisors to conduct orientation properly. Human resource _ _ _ _ _ _ _ **ionals** can provide _ _ **e** new hire _ _ **th** organizational information, _ _ **t** it is _ _ **e** supervisor who _ _ **st** successfully integrate _ _ **e** employee into _ _ **e** work setting.

(Wonghiransombat, 1998, p. 56)

In addition to the construction of new language tests, language teachers should further investigate the students' test-taking strategies in order to validate the language test and to examine what language abilities the test can measure (Cohen, 1994, 1998). Test-taking strategies can be defined as "the processes that the test takers make use of in order to produce acceptable answers to questions and tasks, as well as the perceptions that they have about these questions and tasks before, during, and after responding to them" (Cohen, 1998, p. 216). For instance, some students read an entire cloze passage before filling in the missing parts (Cohen, 1998). Moreover, the perceptions of language tests and test-taking strategies of the students with high- or low-language-ability are different (Cohen, 1984; Sasaki, 2000; Yamashita, 2003). As far as the present researcher has been able to determine, there has been no investigation in Thailand on cloze test-taking strategies. Therefore, cloze completion processes are also included in this study to examine the strategies used in taking the C-Test, the MC-Test, the NC-Test, and the NMC-Test for non-native undergraduate students.

In conclusion, this research is aimed at comparing the new cloze formats (the NMC-Test and the NC-Test) with the older cloze formats (the MC-Test and the C-Test) and to examine the similarities and the differences in these four tests for Thai undergraduate students. Also, this study focuses on examining what test-taking strategies or procedures the students use while responding to the different types of cloze tests.

Copyright by Mahidol University

1.2 Purpose of the Study

The present study aims to investigate the differences in the four types of the cloze tests by comparing the use of the MC-Test with that of NMC-Test, and the use of the C-Test with that of the NC-Test. In order to understand the cloze test-taking strategies, the study is also designed to find out to what extent undergraduate students use seven test-taking strategies while answering the different types of cloze tests. The strategies are based on the latest categorization of Sasaki (2000). The new cloze tests including the C-Test, the NC-Test, the MC-Test, and the NMC-Test were taken by first-year science students at Mahidol University in the first semester of academic year 2003. Therefore, the research questions are posed as follows:

(1) Does the NMC-Test yield different results from the original MC-Test in measuring students' language proficiency?

(2) Does the NC-Test yield different results from the original C-Test in measuring students' language proficiency?

(3) Does using every third-word deletion of the NMC-Test and the NC-Test affect the discrimination power of the test?

(4) What test-taking strategies do the first-year undergraduate students in the Faculty of Science at Mahidol University use while taking the C-Test, the MC-Test, the NMC-Test, and the NC-Test?

1.3 Significance of the Study

This study is designed to compare the new cloze formats, including the original C-Test with the NC-Test, and the original MC-Test with the NMC-Test. The results of this study may provide an alternative way for language teachers to measure the English language proficiency of Thai undergraduate students learning EFL. Test-taking strategies are also studied to enable the language teachers to understand how effectively the students respond to the new types of cloze passage.

Copyright by Mahidol University

1.4 Scope and Limitation of the Study

(1) The study is limited to first year science students at Mahidol University in the first semester of the academic year 2003. The results cannot be generalized to other students, at other university levels, and in other areas.

(2) The study focuses on first-year science students with high- and low-language-ability based on the English Entrance Examination scores, which were reported by the coordinator of the science program.

(3) Only exact word scoring is employed in this study.

(4) It is assumed that all of the first-year science students have had some background knowledge of English up to Mathayomsuksa Six.

1.5 Definitions of Terms

Cloze test refers to a test in which the entire words are rationally or randomly deleted and the student is asked to fill in the missing words (Boonsathorn, 1990, 2000; Wonghiransombat, 1998).

C-Test is a test in which the *second part* or the *second half* of every *second word* is deleted and the student's task is to fill in the deleted parts (Boonsathorn, 1990; Klein-Braley, 1985).

New C-Test (NC-Test) is a test in which the *second part* or the *second half* of every *third word* is deleted and the student is required to fill in the missing parts (Thongsa-nga, 1998).

New Modified C-Test (NMC-Test) is a test in which the *first part* or the *first half* of every *third word* is deleted and the student's task is to fill in the missing parts (Wonghiransombat, 1998).

Modified C-Test (MC-Test) is a test in which the *first part* or the *first half* of every *second word* is deleted and the student is required to fill in the deleted parts (Boonsathorn, 1990, 2000; Wonghiransombat, 1998).

Copyright by Mahidol University

Readability refers to "how easily written materials can be read and understood. Readability depends on many factors, including (a) the average length of sentences in a passage, (b) the number of new words a passage contains, and (c) the grammatical complexity of the language used. Procedures used for measuring readability are known as readability formulae" (Richards, Platt, & Platt, 1993, p. 306).

Test-taking strategies are "the processes that the test takers make use of in order to produce acceptable answers to questions and tasks, as well as the perceptions that they have about these questions and tasks before, during, and after responding to the test" (Cohen, 1998, p. 216).

Copyright by Mahidol University

CHAPTER II
LITERATURE REVIEW AND RELATED RESEARCH

Four types of cloze tests used in the present study were investigated by comparing the original C-Test with the New C-Test (the NC-Test), and the original Modified C-Test (the MC-Test) with the New Modified C-Test (the NMC-Test) in measuring the English-language proficiency of EFL university students. The test-taking strategies obtained from group interviews were also examined so as to see how these test-takers at different language proficiency levels responded to each type of cloze test. In this chapter, further details of theoretical issues and related research articles are reviewed to give a clear understanding of the present study.

2.1 Background of the Cloze Test

The cloze test, initiated by Taylor (1953, cited in Anderson, 1976), is a kind of integrative test in which the entire word is rationally or randomly deleted. The word "cloze" was derived from "closure" in Gestalt psychology, indicating that humans are able to fill in what is missing by using their prior knowledge or their experience (Heaton, 1975; Oller, 1979; Sawatdirakpong, 1980). The students' language proficiency can then be determined by measuring how accurately the students can complete the deleted part of its original passage (Hughes, 2003; Spolsky, 1973). For language teachers, the cloze tests have been widely used in order to determine reading ability and English-language proficiency, since the cloze test is more economical, easier to score, and less time-consuming (McNamara, 2000; Oller, 1979).

In the construction of cloze tests, language teachers should select a suitable passage which responds to the target students' language abilities and the goals of the language tests. Oller (1979) clearly emphasizes that the most appropriate cloze passage should depend on the students' ability level in a class if the purpose of the language teachers is to measure English-language proficiency. The following are guidelines on how to choose appropriate texts for constructing cloze tests: (1) the selected passage should not

Copyright by Mahidol University

contain any bias, such as religion or politics; (2) the selected passage should not contain any specialized terms or content; (3) the difficulty level of the selected passage should be suitable for the target students; (4) the length of the selected passage should be sufficient for the number of items; and (5) the selected text should be complete in itself (Oller, 1979; Raatz & Klein-Braley, 2003).

For text difficulty level, many researchers and practitioners in language testing (Crawley & Mountain, 1995; Leu & Kinzer, 1995; Vacca & Vacca, 2003) recommend that a checklist or a readability formula should be employed to estimate whether the selected passage is suitable for the student's reading level. Fry's readability graph was proposed by Fry (1977) to estimate the readability of a reading text by using the average word length and sentence length of three sample 100-word passages from the selected passages (see Appendix A). Fry's readability graph has been widely used due to the fact that it provides a wide range of reading grade levels from the first grade to the seventeenth grade levels (Anderson, 1983; Duffelmeyer, 1983; Fry, 1977, 1989, 1990, 2002; Leu & Kinzer, 1995; Saetung, 1984; Standal & Betza, 1990). Therefore, Fry's readability graph enables the language teacher to adjust and develop texts before constructing and administering tests for the target students.

Regarding the deletion procedure, there are two deletion methods used in the cloze test: the systematic deletion (the fixed-ratio deletion) and the unsystematic deletion (the rational deletion) (Bachman, 1985; Chapelle & Abraham, 1990; Cohen, 1980; Klein-Braley, 1997; Oller, 1979). The former method refers to every n^{th} word deletion which is suitable for assessing general language abilities because "all classes and types of words have an equal chance of being deleted" (Steinman, 2002, p. 293). The latter method is specific word deletion which is appropriate for a particular purpose, such as testing prepositions. Example 6 is presented in order to provide a clear understanding of both deletion procedures. In addition, Bachman (1985) found that both deletion techniques were equally reliable, although systematic deletion was more difficult than the unsystematic deletion. Nonetheless, different deletion rates affect the validity and the measurement of the cloze test (Alderson, 1979, 1980, 1983, 2000; Jafapur, 1995). For example, changing the rate of deletion in the cloze test makes it measure different language abilities (Jafapur, 1995; Weir, 1990).

Copyright by Mahidol University

Example 6: Unsystematic deletion and systematic deletion

UNSYSTEMATIC DELETION	SYSTEMATIC DELETION
Fill in each blank with a/an/the or no article. People today are quite astonished by _____ rapid improvements in medicine. Doctors are becoming more specialized, and new drugs are appearing on _____ market daily. At _____ same time, _____ people are dismayed by _____ inaccessibility of doctors when they are needed. (Adapted from Cohen, 1994, p. 234)	**Fill in the missing word.** People today are quite astonished by the rapid improvements in medicine. Doctors _____ becoming more specialized, and _____ drugs are appearing on the _____ daily. At the same time, _____ are dismayed by the inaccessibility _____ doctors when they are needed. (Cohen, 1994, p. 234)

Many factors including the deletion procedures, the text length, the number of the items, and the goal of the language test have influence on the forms of the cloze tests. Oller (1979) suggests that the cloze test should generally provide 50 deleted items with a minimum length of 250 words for the passage. However, language teachers are sometimes confused about what the difference between a gap-filling test and a cloze test are. The difference between these two tests is focused on the same point, e.g. the word deleted in each sentence. In a gap-filling test, the deleted item is provided within *one sentence*, whereas the deleted word in the cloze test is given in *a paragraph* or *a passage* (Bailey, 1998). In addition, the gap-filling test is suitable for assessing specific language ability, such as grammar or vocabulary, while the cloze test can measure language proficiency of the target students (Alderson, 2000). The following are six forms of cloze tests.

(1) The fixed-ratio cloze (the random cloze): Every n^{th} word is deleted to be suitable for assessing overall language abilities (Alderson, 2000; Bachman, 1985; Oller, 1979; Steinman, 2002). The following is an example of a sixth-word deletion cloze test.

Copyright by Mahidol University

Example 7: Fixed-ratio cloze test

FIXED-RATIO CLOZE

Fill in the missing words.

People today are quite astonished by the rapid improvements in medicine. Doctors **1)**_____ becoming more specialized, and **2)**_____ drugs are appearing on the **3)**_____ daily. At the same time, **4)**_____ are dismayed by the inaccessibility **5)**_____ doctors when they are needed. **6)**_____ doctors' fees are constantly on **7)** _____ rise, the quality of medical **8)**_____ has reached an abysmal low.

(Adapted from Cohen, 1994, p. 234)

(2) The rational cloze**:** Only specific words are deleted to be appropriate for a particular purpose, such as testing grammar, reading comprehension, and vocabulary (Bachman, 1985). As can be seen in Example 8, a rational cloze test where functional words are deleted to assess grammar is presented.

Example 8: Rational cloze test

RAITONAL CLOZE

Fill in the missing words.

Typically, when trying to test overall understanding **1)** _____ the text, a tester will delete those words **2)** _____ seem to carry the main ideas, or **3)** _____ cohesive devices that make connections **4)** _____ texts, including anaphoric references, connectors, and so **5)** _____. However, the tester then needs **6)** _____ check, having deleted key words, that they **7)** _____ indeed restorable form the remaining context.

(Extracted from Alderson, 2000, p. 210)

(3) The conversational cloze**:** Some words or sentences are deleted to determine the communicative language skills of native- and the non-native-test-takers (Hughes, 2003). The student is required to fill in what is missing in the blanks, as shown in Example 9.

Copyright by Mahidol University

Example 9: Conversational cloze

CONVERSATIONAL CLOZE

David: Hello, Mike. How are you?
Mike: Not too bad, David, and you?
David: O. K. You know, **(1)** _____ been trying to work out **(2)** _____ to go on holiday this year. **(3)** _____ a real problem. I really can't decide where to go. Any ideas?
Mike: Well, I suppose you could try the South **(4)** _____ France.
David: No, I don't really think so. I don't know why, exactly. Maybe it's **(5)** _____ bit expensive down there.

(D. Brown, 1983, p. 159)

(4) The multiple-choice cloze: Every n^{th} word or specific words are deleted and choices of approximately two to five words are also provided for each deleted part. So the multiple-choice cloze (see Example 10), provide more choices, is easier than traditional cloze tests (Chapelle & Abraham, 1990), although its construction seems to be complicated (Hinofotis & Snow, 1980). However, the multiple-choice cloze test can be utilized for testing both specific skills and language proficiency.

Example 10: Multiple-choice cloze

MULTIPLE-CHOICE CLOZE

A farmer's daughter had been out to milk the cows and was returning home, carrying her pail of milk on her head. As she walked along, she **1)** _____ thinking:

(A) started
(B) had to
(C) prepared
(D) began to be

(Extracted from Kaczmarek, 1980, p.152)

Copyright by Mahidol University

(5) The matching cloze: Each deleted word, with or without additional distractors, is usually provided in alphabetical order and put in a column on the right of the cloze passage. This form of language test, featuring ease of construction and scoring, is suitable for measuring specific knowledge of English language such as vocabulary, grammar, and reading comprehension for native and non-native elementary students (Baldauf & Propst, 1979). The students are required to match the correct word provided in the right column with the numbered blanks, as shown in Example 11.

Example 11: Matching cloze

MATCHING CLOZE

Ken and Tom like dogs.
(1) _____ like big brown dogs
(2) _____ little white dogs.
Tom **(3)** _____ a brown dog.
He likes **(4)** _____ with his dog.
His **(5)** _____ is running.
It's going to school with him.

Items
a. and
b. dog
c. has
d. playing
e. they
f. Tom

(adapted from Baldauf & Propst, 1979, p. 323)

(6) The cloze elide: Irrelevant words are added to the original text, and the students' task is to find these additional words and delete them (Alderson, 2000; Steinman, 2002). However, the cloze elide test (see Example 12) is very difficult to construct and is suitable for assessing reading speed (Alderson, 2000).

Copyright by Mahidol University

Example 12: Cloze elide

CLOZE ELIDE

Tests are actually **a** very difficult to construct in this way. One has to be sure **over** that the inserted words do not belong **with**: that it is not possible to interpret **great** the text (albeit in some of different way) with the added words. If so, candidates will not be **therefore** able to identify the insertions.

(Alderson, 2000, p. 226)

The techniques of scoring cloze tests can be divided into two types: exact word scoring and acceptable word scoring. "The exact word method counts only the word deleted from the original passage as correct, whereas the acceptable word method usually counts any contextually acceptable answer as correct" (Brown, 1980, p. 311). Many researchers and practitioners in language testing (Brown, 1980; Lange & Clausing, 1981; Oller, 1972, 1979; Weir, 1990) also support using exact word scoring in the cloze test because it is quick and easy to use; nonetheless, this scoring method seems to be too strict for the test-takers for whom English is not the first language (Oller, 1979). While others (Abraham & Chapelle, 1992; Hinofotis, 1980; Lange & Clausing, 1981; Oller, 1979) state that acceptable word scoring is suitable for measuring the English language proficiency of EFL students, although this scoring technique is more expensive and time-consuming. Therefore, language teachers should take the testing situation and the test purpose into consideration in order to choose the most appropriate scoring method for the language test (Bachman & Palmer, 1996; Brown, 1980, 1996; Hughes, 2003).

Many previous studies of the cloze test indicate that the cloze test is an effective instrument which has great reliability and validity in measuring general language proficiency (Aitken, 1977; Anderson, 1976; J. D. Brown, 1983; Chavanachat, 1986; Fotos, 1991; Jonz, 1987, 1990; McKenna & Layton, 1990; Oller & Conrad, 1971; Oller, 1979; Stubbs & Tucker, 1974; Weir, 1990) and a specific knowledge of target language such as grammar and vocabulary (Alderson, 1979, 1980; Cohen, 1980; Mullen, 1979; Oller & Conrad, 1971; Oller & Inal, 1971). While other studies claim that the cloze test does not measure language ability beyond the sentence level (Alderson, 1979, 1980, 1983; Bachman, 1982; J. D. Brown, 1983; Markham, 1985, 1987; Shanahan et al., 1982). Klein-

Copyright by Mahidol University

Braley (1997) also points out that the construction of the cloze test requires a long passage. So these problems have led to the development of a new form of the cloze test, which is called the C-Test.

2.2 The C-Test vs. the NC-Test

The C-Test, invented by Raatz and Klein-Braley (1981), is a test in which the *second half* of every *second word* is deleted and the student's task is to restore the deleted parts, as shown in Example 13. The original C-Test was constructed as a way of testing English language proficiency besides using cloze tests. However, in the C-Test, the second half of each word must be deleted if the deleted word contains an even number of letters, such as "**m a n y**" (4 letters). For a word with an odd number of letters, its larger part must be deleted, such as **e x c e l l e n t** (9 letters). Many research findings also show that the C-Test is more effective and more reliable than the traditional cloze tests in assessing the students' language proficiency (Babaii & Ansary, 2001; Cohen, Segal, & Weiss, 1984; Connelly, 1997; Dörnyei & Katona, 1992, 1993; Klein-Braley, 1985, 1997) is easy to construct and to score (Babaii & Ansary, 2001; Connelly, 1997; Dörnyei & Katona, 1992, 1993; Weir, 1990).

Example 13: The original C-Test

THE C-TEST

Many foreigners find that Thailand is a very pleasant place to have a holiday. They **disc** _ _ _ _ that **th** _ _ _ are **ma** _ _ interesting **thi** _ _ _ to **d** _ and **t** _ see. **Th** _ _ say **th** _ _ the **bea** _ _ _ _ are **cl** _ _ _ and **t** _ _ scenery **i** _ beautiful. **Ma** _ _ say **th** _ _ the **hot** _ _ _ are **exce** _ _ _ _ _ and **n** _ _ too **expe** _ _ _ _ _. They **exper** _ _ _ _ _ with **diff** _ _ _ _ _ kinds **o** _ Thai **fo** _ _ and **fi** _ _ that **i** _ tastes **deli** _ _ _ _ _.

(Boonsathorn, 1990, p. 48)

Some other studies report problems in using the C-Test in measuring the proficiency in the target language. For example, the C-Test does not assess language abilities beyond the sentence level (Cohen, Segal, & Weiss, 1984; Sigott & Köberl, 1993), and seems to measure the intelligence quotient (IQ) or spelling ability rather than general

Copyright by Mahidol University

language skills (Jafapur, 1995). Some C-Test items, especially the functional words, were reported to have low discrimination power (Cleary, 1988; Jafapur, 1999; Wolter, 2002) and lack of validity (Bradshaw, 1990; Grotjahn, 1986; Jafapur, 1995). Dörnyei and Katona (1992) add that the C-Test is too difficult for EFL secondary students at the secondary level. Consequently, Klein-Braley and Raatz (1984) proposed the following criteria to make the C-Test more reliable and valid in measuring the target language proficiency: (1) the C-Test should contain at least 100 items; (2) the deletion rate and the starting points should be fixed; (3) only exact-word scoring should be employed; (4) the C-Test should contain various passages; (5) native speakers should get a perfect score on the C-Test; and (6) the words affected by the deletion should be a representative sample of the test (p.136).

The previous research on the C-Test focuses on its validation and its measurement of the target language, such as English. In the original study of the C-Test, Raatz & Klein-Braley (1981) examined the use of English and German C-Tests to find out whether the C-Test could be an alternative in assessing the target language. The subjects of the study were divided into two groups. The first group was composed of English native speakers, English-native-speaking schoolchildren, and non-native speakers of English. These students were requested to take the English C-Test. The second group taking the German C-Test consisted of German-native speakers, non-native speakers of German, and German-native-speaking schoolchildren at the third grade level. The results showed that the C-Test had great reliability and validity in assessing the target language of the non-native and the native test-takers.

Dörnyei and Katona (1992) also show that the C-Test is highly reliable and effective in assessing the English language proficiency of Hungarian EFL learners. Their investigation was conducted in order to validate this type of language test for the EFL students. The subjects of the study were 102 Hungarian university students and 53 Hungarian secondary students. These students were then requested to take the C-Test. The results of this study show that the C-Test is suitable to measure language proficiency of non-native students for whom English is not the first language, although this C-Test was reported to be too difficult for Hungarian students at the secondary level.

Copyright by Mahidol University

Connelly (1997) supported using the C-Test to measure the general language proficiency of high-level students studying English as a foreign language. His study examined the English C-Test with non-native postgraduate students studying at the Asian Institute of Technology (AIT) in Bangkok, Thailand. The C-Test with 100-deleted items was administered to EFL postgraduate students from six different countries: Thailand, Vietnam, Taiwan, Indonesia, Japan, and Cambodia. The results of this investigation indicate that the C-Test is highly reliable and has concurrent validity in assessing the language proficiency of English within EFL contexts. However, the C-Test seems to be less effective for EFL students in the lower levels (Cleary, 1988; Connelly, 1997; Dörnyei & Katona, 1992).

This has led to the development of the original C-Test, to make it more suitable for measuring the English language proficiency of non-native students. Thongsa-nga (1998) proposed the New C-Test (the NC-Test) by deleting the *second half* or the *second part* of every *third word* in order to provide more clues for Thai students at the upper secondary level. Thongsa-nga's investigation examined the effect of different starting points in the NC-Tests and students' attitudes towards the measurement using these language tests. In this study, the three forms of the NC-Test, with third, fourth, and fifth starting points were administered to 97 Mathayom Suksa six students at Srakaew School. These participants were also requested to answer a research questionnaire about what skills the NC-Test measured--vocabulary, grammar, reading comprehension, or English language proficiency. Her findings reveal that the NC-Test with the third starting point is the most reliable form for measuring the English language proficiency of Thai Mathayomsuksa six students; nonetheless, the majority of these students reported that the NC-Test seemed to measure vocabulary and reading skills. Thangsa-nga (1998) also adds that the different starting points had an influence on the discrimination power of these three forms of the NC-Tests. Example 14 shows a comparison of C-Test and NC-Test deletion.

Therefore, the present study continues investigating the original C-Test and the NC-Test in order to determine these two language tests are suitable for assessing the general language proficiency of the first-year undergraduate science students studying English as a foreign language.

Copyright by Mahidol University

Example 14: A comparison of word deletion between the original C-Test and the NC-Test

THE C-TEST	THE NC-TEST
Many foreigners find that Thailand is a very pleasant place to have a holiday. They **disc** _ _ _ _ that **th** _ _ _ are **ma** _ _ interesting **thi** _ _ _ to **d** _ and **t** _ see. (Boonsathorn, 1990, p. 48)	There is a dark shadow over schools and colleges where students are now facing the enormous problem of drugs. There seems **t** _ be an **incr** _ _ _ _ in the **u** _ _ of alcohol, **tob** _ _ _ _ and other **dr** _ _ _ by students. (Thongsa-nga, 1998, p. 43)

2.3 The MC-Test vs. the NMC-Test

The Modified C-Test (the MC-Test), also known as the X-Test, was initiated by Boonsathorn (1987). The original MC-Test is a test in which the *first half* of every *second word* is deleted and the students are requested to fill in all the deleted parts, as can be seen in Example 15. In the MC-Test, if the deleted word contains an even number of letters, the first half of this word must be deleted, such as **i n c ome** (6 letters). For a word with an odd number of letters, its larger part must be removed, such as **o b v i ous** (7 letters). In addition, some research findings report that the MC-Test had high reliability and validity in measuring grammatical competence (Prapphal, 1996) and the language proficiency of English for non-native-speaking test-takers (Boonsathorn, 1987; Wonghiransombat, 1998). Nonetheless, Sigott and Köberl (1993) point out that the MC-Test does not measure language abilities beyond the sentence level and seems to be too difficult for EFL test-takers.

Copyright by Mahidol University

Example 15: The original MC-Test

THE MC-TEST

Many foreigners find that Thailand is a very pleasant place to have a holiday. They _ _ _ _ **over** that _ _ _ **re** are _ _ **ny** interesting _ _ _ **ngs** to _ **o** and _ **o** see. _ _ **ey** say _ _ **at** the _ _ _ _ **hes** are _ _ _ **an** and _ _ **e** scenery _ **s** beautiful. _ _ **ny** say _ _ **at** the _ _ _ **els** are _ _ _ _ _ **lent** and _ _ **t** too _ _ _ _ _ **sive**. They _ _ _ _ _ **iment** with _ _ _ _ _ **rent** kinds _ **f** Thai _ _ **od** and _ _ **nd** that _ **t** tastes _ _ _ _ _ **ious**. They _ _ **e** delighted _ _ **th** Thai _ _ _ **ic** and _ _ _ _ _ **nated** by _ _ **ai** dancing. Visitors from all countries often say that Thai people are warm and friendly.

(Boonsathorn, 1990, p. 49)

In the original investigation of the MC-Test, Boonsathorn (1987) compared the MC-Test with the C-Test in measuring language proficiency in English, and examined the reading strategies of ESL students. The subjects of the study included 389 native-speaking-English high school (L1) students and 104 ESL adult learners (L2) in Alberta, Canada. The two types of language tests were administered to both groups. His study showed that the MC-Test and the C-Test were highly reliable for an English test of native and non-native learners even though the MC-Test was reported to be more difficult and had better discrimination than the C-Test. For reading strategies, only 28 ESL adult students were interviewed to report what reading strategies they used while taking these tests. Boonsathorn (1987) added that the ESL learners taking the MC-Test required more strategies than those taking the C-Test.

Other previous studies also support the use of the MC-Test in assessing English language proficiency. Köberl and Sigott (1996) compared the scores of the MC-Test with the scores of the C-Test taken by 82 English native students in United Kingdom and 42 German learners of English, and investigated the "item facilities of these two tests were not influenced by whether the subjects were native and non-native" (p.53). In addition, the results show that the item facilities in the MC-Test and the C-Test highly correlated to both subject groups. For this reason, these two language tests are equally appropriate in measuring language proficiency in English for the native- and the non-native-test-takers.

Copyright by Mahidol University

Prapphal (1996) also constructed two MC-Tests by using General English and Academic English texts in order to find whether the MC-Tests in the study could better measure lexical competence or the grammatical competence. Both the General English MC-Test and the Academic English MC-Test were administered to 48 third-year Thai students in the science program at Chulalongkorn University. The results reveal that a MC-Test constructed from General English or Academic English is highly reliable and has concurrent validity in measuring the grammatical competence; however, Prapphal (1996) claimed that these two MC-Tests seem to measure lexical competence rather than the syntactic competence.

In order to adjust the original MC-Test to be more suitable for measuring the English language proficiency in EFL contexts, Wonghiransombat (1998) then proposed the New Modified C-Test (the NMC-Test) by deleting the *first half* of every *third word*, which provides more clues for non-native-speaking students. A comparison between text deletion in the MC-Test and the NMC-Test is shown in Example 16. The investigation of Wonghiransombat was designed to compare the original MC-Test to the NMC-Tests with different starting points (the NMC-Test with the third starting point, the NMC-Test with the fourth starting point, and the NMC-Test with the fifth starting point). The subjects were 84 postgraduate students studying at the National Institute Development Administration (NIDA) in Bangkok, Thailand. They were requested to take the original MC-Test and one of the three forms of the NMC-Tests. The findings indicate that the NMC-Tests with different deletions were considered to be easier than the original MC-Test and highly reliable as an alternative assessment of the overall language skills of Thai postgraduate students. Wonghiransombat (1998) also reports that the different starting pints did not affect the reliability, the validity, or the difficulty of the three NMC-Tests.

So this present study further examines whether the original MC-Test and the NMC-Test with third-word deletion can be an alternative in measuring the English language proficiency of non-native-speaking tertiary students in the science program.

Copyright by Mahidol University

Example 16: A comparison of word deletion between the original MC-Test and the NMC-Test

MC-TEST	NMC-TEST
As inflation denotes changes in the general price levels pervading the whole economy, a number of distortions occur. One _ _ _ _ **ous** effect _ **f** inflation _ **s** in _ _ _ **ome** and _ _ _ **lth** distribution. (Wonghiransombat, 1998, p. 53)	Acid rain, endangered species, lead poisoning, the destruction of the ozone layer, waste disposal – the list of environmental problems today seems endless. It's a _ _ _ **tty** grim picture; _ _ _ _ **ver**, we must _ _ _ **ept** the challenge; _ _ _ **re** is hope. (Wonghiransombat, 1998, p. 58)

2.4 Test-Taking Strategies

Test-taking strategies can be defined as "the processes that the test takers make use of in order to produce acceptable answers to questions and tasks, as well as the perceptions that they have about these questions and tasks before, during, and after responding to them" (Cohen, 1998, p. 216). Generally, the processes of taking language tests are divided into two types: "the process of responding" and "the reactions to items and subtests" (Cohen, 1984, pp. 71-72). The former involves the strategies that the students use while taking the language tests. For example, some students use context clues to restore the deleted parts in the cloze tests (Babaii & Ansary, 2001). The latter focuses on the test-takers' attitudes towards the language tests. For instance, some students prefer to take the C-Test, which provides a chance of guessing, rather than to take the traditional cloze test (Weir, 1990).

In order to identify the test-taking strategies used by the target respondents, investigations can be done by observation, performance analysis, questionnaires, and interviews (Cohen, 1994). The results of test-taking strategies also enable the language teachers to validate the language test and to determine what language abilities this language test can measure (Cohen, 1994, 1998). Many previous studies of test-taking strategies concentrated on the measurement of cloze tests rather than the completion processes. Some studies (Chàvez-Oller et al., 1985; Fotos, 1991; Jonz, 1987, 1990;

Copyright by Mahidol University

McKenna & Layton, 1990; Oller, 1973; Oller & Conrad, 1971; Sasaki, 2000; Storey, 1997; Yamashita, 2003) found that cloze tests were highly reliable in assessing overall language skills. The students required both syntactic and semantic information in order to fill in the cloze passage (Oller, 1979). Nonetheless, some others (Alderson, 1979, 1983; Bachman, 1982; J. D. Brown, 1983; Markham, 1985, 1987; Shanahan et al., 1982) reported that the cloze tests did not measure language abilities beyond the sentence level, because the students sometimes were able to fill in the cloze items by using only lexical competence (Alderson, 1979, 1983, 2000).

Many researchers and practitioners in language testing (Bachman, 1985; J. D. Brown, 1983; Jonz, 1990; Markham, 1985, 1987; Yamashita, 2003) were aware that cloze tests were sensitive to the text-level constraints. For example, each cloze item did not contain the same information. That depended on the types of the words deleted (Alderson, 1980; Bachman, 1982, 1985; Jonz, 1990; Yamashita, 2003). Some items, such as prepositions and articles, could be restored by using only "linguistic knowledge" while some others, such as anaphora, lexical repetition, and conjunction required "textual understanding"(Yamashita, 2003, p. 268). Therefore, students probably use different strategies based on the type of deleted words. Bachman (1985) developed a framework dividing cloze items into four categories according to the textual information that students use while answering each cloze item: (1) Within Clause; (2) Across Clause, Within Sentence; (3) Across Sentences, Within Text; and (4) Extratextual.

Bachman (1985) studied performance on fixed-ratio and rational cloze tests in order to examine what language skills cloze tests with different deletions measured. These cloze tests were administered to 910 participants (native- and non-native-speaking college students) in Illinois. The results show that these students frequently used the 'Extra-textual' strategy for fixed-ratio cloze tests while the students taking a rational cloze test mostly employed 'Across Sentences, Within Text' strategy. So both fixed-ratio and rational cloze tests have high reliability in assessing the English language proficiency, although all cloze items did not measure the same language abilities. Bachman also suggests that there should be further studies on the test items in the different forms of the cloze tests.

Copyright by Mahidol University

However, other researchers (Sasaki, 2000; Yamashita, 2003) examined test-taking strategies by using verbal reports to see how the students responded to the language tests. Sasaki (2000) studied the effects of cultural schemata on how Japanese EFL students responded to unfamiliar and familiar fixed-ratio cloze passages. The subjects were 60 Japanese EFL students with the same English reading proficiency level. These students were divided into two groups; each group was required to complete culturally familiar or culturally unfamiliar cloze passages (Sasaki, 2000). The students were asked to report their test-taking strategies to be categorized based on the modified cloze test-taking strategies categorization of Bachman (1985). The results show that the students reading the familiar cloze passage could answer more items and used these three categories of test-taking strategies--'Within Clause', 'Across Clause, Within Sentence' and 'Extratextual' more frequently than the students reading the unfamiliar cloze passage (Sasaki, 2000). Sasaki pointed out that the previous studies of cloze test-taking frameworks were conducted based on discourse analysis that focused too much on textual relation, so Items 3, 6, and 7 were added to Bachman's framework (1985, p. 95), as shown in the following:

(1) *Within Clause*: The examinee uses information provided only by the clause in which an item appears.

(2) *Across Clause, Within Sentence*: The examinee uses information provided by a broader context than the clause in which an item appears, but a narrower context than the orthographic sentence.

(3) *Across Sentences, Within Paragraph*: The examinee uses information provided by the broader context of the orthographic paragraph containing an item.

(4) *Across Paragraphs, Within Text*: The examinee uses information provided by the context of the entire text.

(5) *Extratextual*: The examinee uses information that is not provided by the text itself, but which is assumed to be included in the examinees' world knowledge.

(6) *Guessing*: The examinee guesses at the answer.

(7) *Missing*: The examinee does not/cannot say anything about his or her test-taking processes, or does not answer the item.

Copyright by Mahidol University

Sasaki's (2000) framework was also utilized by Yamashita (2003) to examine what test-taking strategies 12 Japanese students (six skilled readers and six less skilled readers) used for a rational deletion cloze test to find whether this type of test could measure "higher order processing ability" (p. 267). Think-aloud protocols were employed by both groups to report how they answered the rational cloze test. In addition, their verbal reports were tabulated, based on Sasaki's cloze test-taking strategies framework (2000). The findings of Yamashita (2003) showed that both skilled and less skilled readers used the 'Across Paragraphs, Within Text' strategy more frequently than other strategies. The skilled readers used Item 3 (Across Sentences, Within Paragraph) and Item 4 (Across Paragraphs, Within Text) more than the less skilled readers (Yamashita, 2003). Yamashita (2003) thus concludes that the rational deletion cloze test could measure English language proficiency of EFL students. Consequently, this present study focuses on the categorization framework of cloze test-taking strategy by Sasaki (2000) to report how the voluntary subjects answered each of the four cloze tests. This process is most likely to show the ways the students answered the four cloze tests used in the study.

In summary, this research is designed to compare the original C-Test with the NC-Test and the original MC-Test with the NMC-Test in measuring the English language proficiency of first-year undergraduate science students at Mahidol University. The results of the study may provide language teachers with several choices of cloze tests to assess general language skills of EFL tertiary students. Test-taking strategies are also investigated so that the language teachers can explore how these target students at high and low language abilities responded to each of the cloze tests.

Copyright by Mahidol University

CHAPTER III
METHODOLOGY

The present study was designed to compare the original C-Test (the C-Test) with the New C-Test (The NC-Test), and the original Modified C-Test (the MC-Test) with the New Modified C-Test (The NMC-Test) in measuring EFL undergraduate students' English language proficiency. The test-taking strategies used by the subjects were also investigated to find how the target students responded to each of the four cloze tests. Further details including the subjects, the research instruments, data collection, and the statistical devices used are presented as follows.

3.1 Subjects

The target students of the present study were 292 first-year undergraduate science students studying in the Faculty of Science at Mahidol Univeristy in the first semester of the academic year 2003. The coordinator of the first-year science students divided these science students into three groups according to language proficiency levels (high, mid, and low), based on the scores of the English Entrance Examination. This study focused merely on the two extreme groups of language proficiency levels: High and Low. There were a total number of 114 students, and yet, at the start of the cloze-test administration, there were two students in each group reported missing. Therefore, the exact number of participants in the study was 110. The undergraduate science students in each English class were then divided into four sub-groups in order to take one of the following cloze tests: the C-Test, the NC-Test, the MC-Test or the NMC-Test. These cloze tests were administered to those 110 participants in an equal proportion at the start of their English classes, as shown in Table 2. Four voluntary students from both the high- and low-language-ability groups were subsequently requested to report what test-taking strategies these individual volunteers used while taking each type of cloze test.

Copyright by Mahidol University

Table 2: Numbers of the high and low groups taking each type of cloze test

Group[a]	Test Type				Total
	C-Test	NC-Test	MC-Test	NMC-Test	
High	14	13	14	14	55
Low	14	13	14	14	55

[a] The total number of the participants in the study was 110.

3.2 Research Instruments

The instruments used in the study were four types of cloze test: the original cloze Test (The C-Test), the New C-Test (the NC-Test), the original Modified C-Test (the MC-Test), and the New Modified C-Test (the NMC-Test). These language tests were administered at the start of the English classes in order to compare the test scores between the C-Test and the NC-Test and between the MC-Test and the NMC-Test taken by the target students. Each cloze test included 100 deleted items and was composed of four short different passages in general sciences: (1) Human disease; (2) Humans and computers: a working relationship; (3) No smoking please; and (4) Try jogging for fitness. For the selection of those four reading texts, the following criteria were established to find suitable passages for EFL undergraduate students: (1) the selected passages in general sciences for EFL university students should range from high-beginner (equivalent to the tenth grade) to low intermediate level (equivalent to the thirteenth grade, or college level) estimated by the publishers and Fry's readability graph (1977, 2002) because the target students have already obtained some background knowledge of English up to Mathayom Suksa Six (equivalent to the twelfth grade); (2) the content of each selected passage must be complete within one paragraph and must not include illustrations or numbers; and (3) all of the selected passages should be expository written texts, each of which contains 25 items of second or third-word deletion.

The researcher selected scientific texts from various sources whose text difficulty levels was subsequently estimated by Fry's readability graph (1977, 2002). Additionally, the difficulty levels of the selected passages were also related to two main factors: vocabulary and sentence complexity (Anderson, 1983; Fry, 1977, 1989, 1990, 2002; Leu & Kinzer, 1995). As a consequence, if a passage contains longer words and sentences,

Copyright by Mahidol University

then the passage will be more difficult to understand (Standal & Betza, 1990). Fry's readability graph has been widely used due to the fact that it provides a wide range of reading grade levels from the first grade to the seventeenth grade levels (Anderson, 1983; Duffelmeyer, 1983; Fry, 1977, 1989, 1990, 2002; Leu & Kinzer, 1995; Saetung, 1984; Standal & Betza, 1990). The estimation of Fry's readability level can be done by randomly selecting three 100-word passages each from the beginning, the middle and the end of a book or a fairly long text, then counting the average numbers of syllables and sentences, and then plotting the average numbers of syllables and sentences on Fry's readability graph to show an approximate grade level for each selected reading text. Example 15 shows how to count the number of syllables and sentences in the three selected reading passages. Table 3 displays the average number of syllables and sentences and the plotted place on Fry's readability graph.

Example 17: The counting method of Fry's readability graph for the book entitled "Timed Reading Plus: 25 Two-Part Lessons with Questions for Building Reading Speed and Comprehension Book 5" (Spargo, 2002b)

The first 100 words extracted from the beginning section of the book

/ / / /// / / /// /// / // // / / / // /
It was a beautiful day in Chicago. Catherine was almost skipping as she walked beside her
// / / / / / / //// //// /// / //
parents through the gate of the World's Columbian Exposition. Chicago has campaigned
/ / / // / // / // // / /// ///// / ///
hard for the honor of holding the event marking the 400th anniversary of Christopher
/// / // / //// / // // / / / // / ///
Columbus's first visit to America, just barely edging out New York City. As Catherine
/ / // // / // / / / // / // / / // /
and her parents entered the fairgrounds, they saw a dazzling "White City" in the style of
// //// //// / // // / / /// /
classic European architecture. The city's columns, domes, arches, staircases, and
// / // / / // / // / / / / // ///
fountains were painted in a gleaming white. Inside was a world of future. Catherine
// /// / /// **100 word**
stared openmouthed in amazement ↑ at massive machines, an electric stove, product

displays, and art from all 48 states and other nations as well (p. 27).

Copyright by Mahidol University

The second 100 words extracted from the middle section of the book

Every infectious disease has an incubation period. This is the length of time between the pathogen's gaining a foothold in the body and the appearance of the first symptoms of the disease. Several factors also determine whether a person will become the victim of a disease after being infected. The number of invading germs in the dose of the infection influences the outbreak of disease. So does the virulence of the pathogens; that is their power to do harm. In addition, the condition of the body's immunological defenses also affects the probability of catching a disease. A great many infectious ↑ diseases are

100 word

contagious; that is, they can easily be passed between people (p. 77).

The third 100 words from the ending section of the book

From prehistoric times to the present, there have been many mass migrations of people throughout the world. In a few isolated locations, however, certain tribal or ethnic groups have lived without migrating for many thousands of years. Such people are called aborigines, from Latin, meaning "from the beginning". Aboriginal people lived in areas remote from other cultures, and their existence became known to the rest of the world only when outsiders intruded upon their territories. Some anthropologists in the twentieth century question whether aborigines have always lived in the locations where they have been found in modern times. It is ↑ possible that some aborigines did migrate, but in a

100 word

period so remote in time that there is no record of their migration (p. 111).

Copyright by Mahidol University

Table3: The average number of sentences and syllables of the three selected reading passages shown in Example 15 and the plotted place on Fry's readability graph

	Syllables	Sentences
1st 100 words in the beginning	170	6.21
2nd 100 words in the middle	169	6.25
3rd 100 words in the end	172	5.12
Totals	511	17.5
Average Number [a]	**170.33**	**5.86**

[a] The approximate grade level was the 13th grade.

The approximate grade level based on Fry's readability graph (1977, 2002)

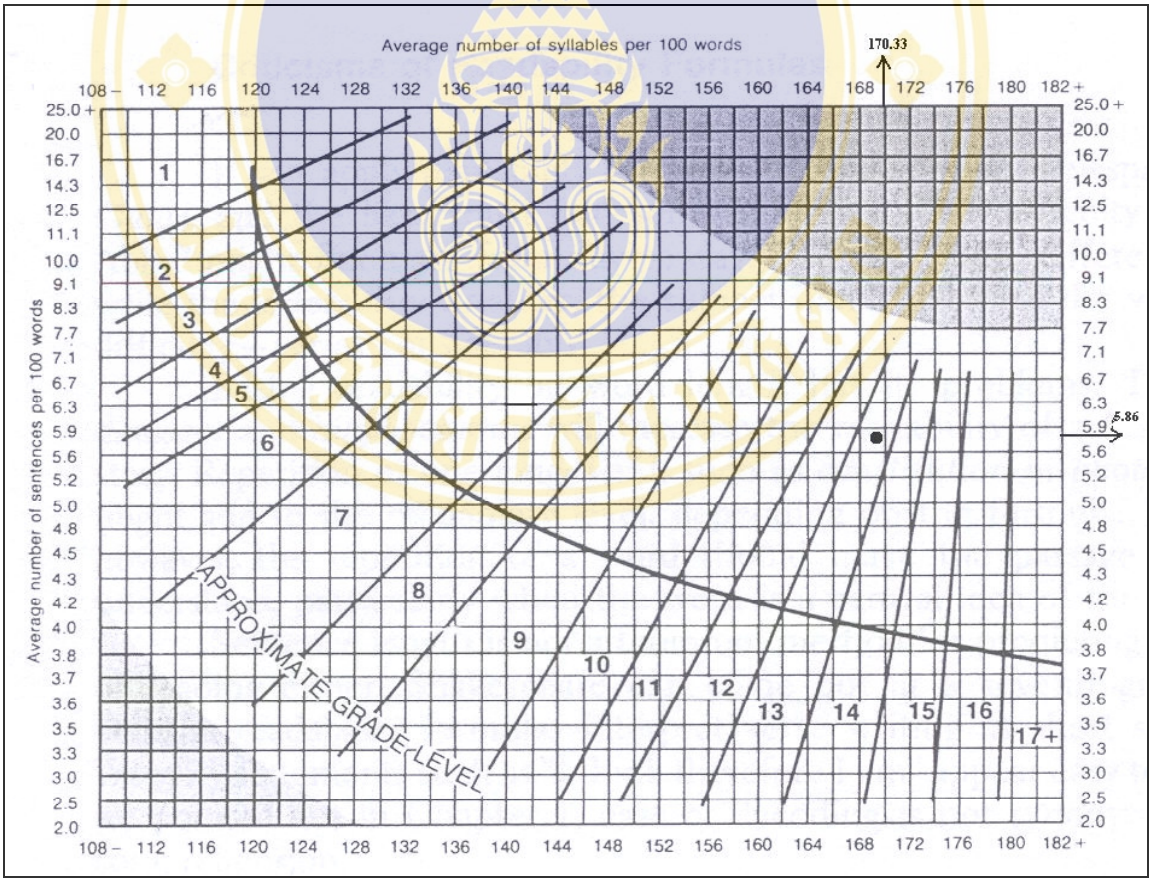


Figure 1: The grade level estimated by Fry's readability graph (1977, 2002)

Copyright by Mahidol University

For the construction of the cloze tests, the researcher selected twelve passages and then had a meeting with two subject teachers in the first-year science program in order to have them confirm that these selected passages were suitable for the reading levels of the target students. Six out of the total number of twelve passages were then chosen for the pilot science students. As a result, the difficulty levels of the selected passages used in the present study were verified in three ways: by the publishers, the subject teachers, and Fry's readability graph (1977, 2002). Due to the piloting of the constructed cloze tests, two passages had to be removed because they seemed to be too easy for the target students (see Section 3.3). Therefore, four short passages ranging from the tenth grade level to the thirteenth grade, or college level, estimated by Fry's readability graph, as can be seen in Table 4, were ultimately used to construct four types of cloze test in this study.

Table 4: Sources and estimated readability levels of the selected passages used in the present study

Topics	Source	Words	Readability Levels
Human Disease	*Timed Reading Plus: 25 Two-Part Lessons with Questions for Building Reading Speed and Comprehension Book 5* by Spargo, E. (Jamestown) 2002: 77	122	13th grade
People and Computers: a working relationship	*Past, Present and Future: a Reading-Writing Text by* Gregg, J. Y. & Russell, J. (Wadsworth Publication) 1990:103-106	125	12th grade
No Smoking, Please	*Concepts and Comments: a Reader for Students of English As a Second Language* by Ackart, P. (Harcourt Brace) 1997: 298-300	120	11th grade
Try Jogging for Fitness	*Timed Reading Plus: 25 Two-Part Lessons with Questions for Building Reading Speed and Comprehension Book 4* by Spargo, E. (Jamestown) 2002: 73	115	10th grade

Copyright by Mahidol University

3.2.1 The Construction of the C-Test

The C-Test was first developed by Raatz and Klein-Braley in 1981 to measure the foreign language proficiency besides the traditional cloze tests. The construction of the C-Test was also known as "the rule of 2" (Klein-Braley & Raatz, 1984, p. 136). That is the *second half* or *second part* of every *second word* is deleted as can be seen in Example 16. Many studies (Connelly, 1997; Dörnyei & Katona, 1992; Klein-Braley, 1985, 1997; Klein-Braley & Raatz, 1984; Raatz & Klein-Braley, 1981, 2003) found that the C-Test was highly reliable in measuring language proficiency of such target languages as English and provided more test items within shorter texts (Klein-Braley, 1997). Other studies pointed out that the C-Test was too difficult for the non-native-speaking secondary students (Dörnyei & Katona, 1992), and seemed to measure IQ or spelling ability rather than English language proficiency (Jafapur, 1995). Therefore, Klein-Braley and Raatz (1981) established the following criteria to construct a C-Test more effectively: containing at least one hundred items, scoring exact words, and fixing the deletion and the starting points. For the present study, the C-Test was designed to compare whether the original C-Test differed significantly from the NC-Test in measuring the English language proficiency of the undergraduate science students. The construction of the C-Test also used the following procedures:

(1) Delete the *second half* or *second part* of every *second word* (see Example 18).

(2) Divide every second word into two parts. For a word with an odd number of letters, such as "**cause**", containing five letters, its larger part, comprising three letters, **ca u s e,** must be deleted, while the first half contains only two letters**.** For a word with an even number of letters, such as "**follow**", with six letters, each part contains three letters, so the second half, **fol l o w**, must be deleted (see Appendix B).

Copyright by Mahidol University

Example 18: The C-Test

THE C–TEST

Pesticides can help control pests. But **the y** can **cau s e** harm, **to o**. People **ma y** not **foll o w** the **direct i o n s** carefully **whe n** using **the m**.

3.2.2 The Construction of the New C-Test

The NC-Test, imitating the construction of the C-Test, is a test in which the *second half* or *second part* of every *third word* is deleted and the student's task is to fill in the deleted parts, as can be seen in Example 19. The NC-Test was initiated by Thongsa-nga in 1998 as an instrument testing English language proficiency of Thai Mathayomsuksa Six students. Her study showed that the NC-Test with third-word deletion was more reliable and appropriate for the Mathayom Suksa Six students than the original MC-Test; however, most respondents thought that the NC-Test measured their vocabulary skills rather than their language proficiency in English. Thongsa-nga (1998) then recommended further study of the use of the NC-Test to validate the use of this language test for non-native students for whom English was not the first language. So the NC-Test with third-word deletion in this study was constructed in order to provide more clues for first-year undergraduate science students and to find whether there was any significantly difference between the NC-Test and the original C-Test in testing language proficiency in English. The construction of the NC-Test is described below.

(1) Delete the *second half* or *second part* of every *third word* (see Example 19).

(2) Separate every third word into two parts. For a word with an odd number of letters, such as "**using**", containing five letters, its larger part comprising three letters, **us i n g,** must be deleted, while the first half contains only two letters**.** For a word with an even number of letters, such as "**directions**", with ten letters, each part contains five letters, so the second half, **direc t i o n s**, must be deleted (see Appendix C).

Copyright by Mahidol University

Example 19: The NC-Test

THE NC-TEST

Pesticides can help control pests. But they **ca n** cause harm, **to o**. People may **no t** follow the **direct i o n s** carefully when **usi n g** them.

3.2.3 The Construction of the Modified C-Test

The Modified C-Test (the MC-Test) or the X-Test, initiated by Boonsathorn (1987) is a test in which the *first half* or the *first part* of every *second word* is deleted, as shown in Example 20. The subjects were requested to fill in the deleted parts. Additionally, some research findings (Boonsathorn, 1987, 1990; Prapphal, 1994, 1996; Wonghiransombat, 1998) showed that the MC-Test had great reliability and validity in measuring the language proficiency of English for high-level EFL students; nonetheless, the MC-Test seemed too difficult for the native-speaking test-takers and did not measure language abilities beyond the sentence level (Sigott & Köberl, 1993). So, the MC-Test in this study was constructed in order to find out whether the MC-Test could be used to measure the language proficiency of the non-native-speaking university students in the science program, and to compare whether the original MC-Test yielded any different results from the simplified one, the NMC-Test. The construction of the MC-Test used the following procedures:

(1) Delete the *first half* or *the first part* of every *second word* (see Example 20).

(2) Divide every second word into two parts. For a word with an odd number of letters, such as "**may**", containing three letters, its larger part, comprising two letters, **m a y**, must be deleted, while the second half contains only one letter**.** For a word with an even number of letters, such as "**follow**" with six letters, each part contains three letters, so the first half, **f o l low**, must be deleted (see Appendix D).

Copyright by Mahidol University

Example 20: The MC-Test

THE MC-TEST

Pesticides can help control pests. But **t hey** can **c a use** harm, **t oo**. People **m ay** not **f o llow** the **d i r e ctions** carefully **w hen** using **t hem**.

3.2.4 The Construction of the New Modified C-Test

The New Modified C-Test (the NMC-Test) was initiated by Wonghiransombat (1998) as an alternative to measure general language skills instead of the original MC-Test which seems too difficult for native-speaking students (Sigott & Köberl, 1993). Wonghiransombat (1998) then constructed the NMC-Test by deleting the *first half* of every *third word* to provide more clues for EFL test-takers. Her findings showed that the NMC-Test with third-word deletion was easier than the original MC-Test and more appropriate for Thai postgraduate students. She also added that further investigation of the NMC-Test should be conducted to validate this type of language test for the non-native-speaking students. So, the NMC-Test in the present study was designed to see whether the NMC-Test yielded any significantly different results from the original MC-Test in measuring the English language proficiency of first year undergraduate science students at high and low language ability levels. The NMC-Test was constructed by imitating the original MC-Test, as can be seen in the following.

(1) Delete the *first half* or *the first part* of every *third word* (see Example 21).

(2) Separate every third word into two parts. For a word with an odd number of letters, such as "**too**", containing three letters, its larger part, comprising two letters, **t o o,** must be deleted, while the second half contains only one letter**.** For a word with an even number of letters, such as "**directions**" with ten letters, each part contains five letters, so the first half, **d i r e c tions**, must be deleted (see Appendix E).

Copyright by Mahidol University

Example 21: The NMC-Test

THE NMC-TEST

Pesticides can help control pests. But they **c an** cause harm, **t o o**. People may **n ot** follow the **d i r e ctions** carefully when **u s ing** them.

3.3 Scoring Procedure

Exact word scoring was used with the four forms of the cloze tests: one point each for a correct item and zero points each for a wrong item. In this study, the total score of each newly constructed cloze test was 100. The scores of these four cloze tests were checked and verified by two inter-raters and working teachers of EFL at the tertiary level.

3.4 Piloting Four Different Forms of Cloze Tests

Four different types of cloze tests, including the C-Test, the NC-Test, the MC-test, and the NMC-Test, were piloted with 16 first-year science students at Mahidol University in the academic year 2002. The main purpose was to see if these cloze tests were too difficult or too easy for the target students in the first-year science program. The pilot study was also helpful in investigating how long the test duration should be. Although each cloze test may be done within 60 minutes, some pilot students were allowed to complete the missing parts with no time limit, because they were considered as low-language-ability students. Originally, each cloze test consisted of six short passages including 150 deleted items (6 passages x 25 items = 150 deleted items). These cloze tests were examined to determine which passages were too easy or too difficult for the target students. For example, if most pilot students got full marks from any of the six passages, that selected passage was considered to be too easy and had to be removed. If most pilot students got scores of 50 percent or less, that passage was considered to be too difficult and had to be removed as well.

Copyright by Mahidol University

These cloze tests were administered to these 16 pilot students during the first week of their English classes in the first semester. The pilot students were divided into four sub-groups in equal proportion to take the C-Test, the MC-Test, the NC-Test, or the NMC-Test. In addition, the pilot students were asked to fill in all of the deleted parts. According to the results of the pilot study, the pilot students got full marks for two passages, indicating that these two cloze passages were too easy for the first-year science students. So, they had to be removed from this investigation. The cloze tests used in the present study ultimately consisted of four passages including 100 deleted items (4 passages x 25 items = 100 deleted items). The actual cloze tests used in the study were administered to first-year students studying in the Faculty of Science at Mahidol University at the start of their English classes in the first semester of the academic year 2003. The scores of the newly constructed cloze tests in the pilot study are presented in Appendix F.

The reliability coefficient and the consistency of the language proficiency tests, can be estimated by using the Kuder-Richardson formulas (Bachman, 1990; Brown, 1996; Hughes, 2003; Lyman, 1998). In this study, The Kuder-Richardson formula 21 (K-R 21) was utilized to determine the reliability of these cloze tests in measuring the English language proficiency of the first-year undergraduate science students at Mahidol University. The reason for using this formula to determine the reliability of the non-native-speaking students was that it was quick to calculate (Bachman, 1990; Brown, 1996) even though the K-R 21 might underestimate the reliability coefficient (Brown, 1996). The acceptable reliability coefficient of the proficiency tests should be .80 or higher to identify good assessment (Bachman, 1990; Cohen, 1994; Lyman, 1999). The mean scores and the standard deviations of the new cloze tests were then calculated, tabulated, and analyzed (see Appendix F). Table 5 shows the reliability coefficient of these four cloze tests used in the present study.

Copyright by Mahidol University

Table 5: The reliability of the four new forms of the cloze tests, using the K-R 21

Type	Means	S. D.	Reliability
The C-Test	80.75	23.16	.9648
The NC-Test	69.50	19.62	.9470
The MC-Test	69.50	17.60	.9285
The NMC-Test	73.50	19.09	.9417

3.5 Interview

As far as this study is concerned, there has been no previous investigation on cloze test-taking strategies done in Thailand. So the present study was designed to investigate what strategies the EFL undergraduate science students used while answering the four types of cloze test: the C-Test, the NC-Test, the MC-Test, and the NMC-Test. In the study of Babaii and Ansary (2001), some students use a grammatical knowledge and context clues while filling in the deleted part in a cloze passage. Therefore, the interview data on the use of test-taking strategies reported by the target students could help the researcher to determine how the first-year undergraduate science students at high and low language ability levels responded to the four types of the cloze tests.

The investigation of test-taking strategies can be done by questionnaires, interviews, observation schedules, performance analysis, or verbal protocols (Cohen, 1994). In this study, an interview was used to collect the test-taking strategies of these undergraduate science students. Generally, the interview can be divided in two types: an individual interview and a group interview. The former is suitable for investigating individual performance (O'Malley & Chamot, 1990) although it takes time consuming (O'Malley & Chamot, 1990). The latter can be used to examine the strategies within a small group of three to five students (O'Malley & Chamot, 1990). Moreover, the group interview creates an intimate atmosphere and encourages brainstorming by providing a wide range of responses (G. Anderson, 1998; Cohen, Manion & Morrison, 2000) although it might be "influenced by what other students have said" (O'Malley & Chamot, 1990, p. 95). Therefore, a group interview was employed in this study to find out how the first-year science students responded to the four types of cloze test used in this study.

Copyright by Mahidol University

Semi-structured interview questions were then constructed carefully, based on the investigation of Cohen (1984), and were piloted to see whether the interview questions were useful in identifying the test-taking strategies of the target respondents. In addition, these interview questions were divided into two parts. The first part dealt with the students' backgrounds, while the second part focused on the students' cloze test-taking strategies and the attitudes towards the use of the new cloze tests in measuring the English language proficiency. Four questions about cloze test-taking strategies were constructed as follows: (1) Could you please rate the number of difficulty level of each passage (1 is for the easiest → 4 is for the most difficult)? (2) What procedures do you use while answering the new forms of the cloze tests? (3) Do you read from the beginning to the end, or only parts of the test, or jump around?, and (4) Do you use other ways to find the answer in the cloze tests?

These interview questions were adjusted before using them with the subjects at high and low language proficiency levels to examine the use of cloze test-taking strategies in the MC-Test, the NMC-Test, the C-Test, and the NC-Test. Each group interview took approximately 30-40 minutes. The students' responses were collected by note-taking and tape-recording with their permission. After that, the interview data were transcribed and tabulated to see what categories the students used to fill in the deleted parts. The following is the categorization framework of cloze test-taking strategies proposed by Sasaki (2000) and employed in the present study (see Table 6).

Copyright by Mahidol University

Table 6: Cloze test-taking strategies categorization by Sasaki (2000, p. 95)

The Categorization Framework of Cloze Test-Taking Strategies
(1) ***Within Clause*****:** The examinee uses information provided only by the clause in which an item appears.
(2) ***Across Clause, Within Sentence*****:** The examinee uses information provided by a broader context than the clause in which an item appears, but a narrower context than the orthographic sentence.
(3) ***Across Sentences, Within Paragraph***: The examinee uses information provided by the broader context of the orthographic paragraph containing an item.
(4) ***Across Paragraphs, Within Text*****:** The examinee uses information provided by the context of the entire text.
(5) ***Extratextual*****:** The examinee uses information that is not provided by the text itself, but which is assumed to include the examinees' world knowledge.
(6) ***Guessing*****:** The examinee guesses at the answer.
(7) ***Missing*****:** The examinee does not/cannot say anything about his or her test-taking processes, or does not answer the item.

3.6 Data Collection

High and low language ability groups of first-year undergraduate science students were requested to fill in all the deleted items in the C-Test, the NC-Test, the MC-Test, and the NMC-Test. These four cloze tests were simultaneously administered at the start of the target students' English classes. The tests lasted approximately 60 minutes. After that, the students' answers to each cloze test were checked by two inter-raters: a lecturer of EFL at the tertiary level and the researcher, a part-time English teacher at the university level. The results of the cloze tests scores were then calculated using SPSS for Window Version 11.5 (2002). The statistical devices used in the study were frequency counts, mean scores, standard deviation, the Independent Samples t-test, and Kuder-Richardson Formula 21 in order to find whether there were any significant differences between the students' tests scores on the C-Test and those on the NC-Test, and between the students' tests scores on the MC-Test and those on the NMC-Test.

For cloze test-taking strategies, the volunteer students were interviewed in two small groups: high and low to examine what test-taking strategies these volunteer students

Copyright by Mahidol University

used while taking each of the four cloze tests. Each volunteer group consisted of four students who took the C-Test, the NC-Test, the MC-Test, or the NMC-Test. During the interview, the cloze-test papers were given to each student in both the high and low groups to help remind them of test-taking strategies that they had used while taking these cloze tests. Each group interview took approximately 30-40 minutes. After that, the students' responses were transcribed and analyzed based on Sasaki's (2000) categorization framework of cloze test-taking strategies.

3.7 Data Analysis

The following statistical devices were used to analyze the results of each cloze test, the third-word deletion techniques, and the interview data according to the research questions: (1) Does the NMC-Test yield different results from the original MC-Test in measuring students' language proficiency? (2) Does the NC-Test yield different results from the original C-Test in measuring students' language proficiency? (3) Does using every third-word deletion of the NMC-Test and the NC-Test affect the discrimination power of the test?, and (4) What test-taking strategies do the first-year undergraduate students in the Faculty of Science at Mahidol University use while taking the C-Test, the MC-Test, the NMC-Test, and the NC-Test?

For the first and the second research questions, *the Means and the Standard Deviation* were used as the primary tool to identify the average scores for the four types of the cloze test in measuring the students' English language proficiency levels. These results revealed on which type of cloze test the subjects achieved higher scores. Moreover, the *independent samples t-test* was employed to examine whether there were any significant differences between the students' tests scores on the original C-Test and the NC-test, and between the original MC-Test and the NMC-Test. In addition, the reliability coefficient of each cloze test was estimated by *Kuder-Richardson Formula 21 (K-R 21)*. For good assessment, the acceptable reliability of the proficiency tests should be equal to .80 or greater (Bachman, 1990; Cohen, 1994; Lyman, 1998). The K-R 21 was quick to calculate (Bachman, 1990; Brown, 1996), even though the K-R 21 might underestimate the reliability coefficient (Brown, 1996). The reliability coefficient helped the researcher to determine which cloze tests used in this study were reliable and suitable for these Thai undergraduate science students for whom English is not the first language.

Copyright by Mahidol University

Regarding the third research question, *Item Analysis* was used to analyze all deleted items in the four types of the cloze test. Item analysis consists of two calculations: "*Item Facility (IF)*" showing the difficulty level of the deleted items in these four cloze tests and "*Item Discrimination (ID)*" indicating how well each cloze item distinguishes the high language proficiency group from the low one. Additionally, the acceptable facility values should range from .20 to .80 and the discrimination indexes should be equal to .30 or greater (Nuttal & Skurnik, 1969; Wongsothorn, 1996). So the results of the item analysis were examined to find out whether the different starting points or the third-word deletion in the NMC-Test and the NC-Test had an effect on test discrimination power.

To answer the fourth research question, the students' responses obtained from the group interviews were transcribed and categorized based on the latest category of seven cloze test-taking strategies developed by Sasaki (2000). The interview data were analyzed by an expert and the researcher to find out how the respondents at high and low language proficiency levels responded to the C-Test, NC-Test, the MC-Test, and the NMC-Test. In addition, the results of the cloze test-taking strategies were then reported by using *Frequencies*. *Pearson's Product Moment* was also used in this study to examine the correlation of the test-taking strategies indicated by two inter-raters.

Copyright by Mahidol University

CHAPTER IV
RESULTS

In this chapter, the scores on the four cloze tests, the results of using third-word deletion, and the interview data obtained from the first-year undergraduate science students with high and low language abilities were further investigated to answer the following research questions: (1) Does the NMC-Test yield different results from the original MC-Test in measuring students' language proficiency? (2) Does the NC-Test yield different results from the original C-Test in measuring students' language proficiency? (3) Does using every third-word deletion of the NC-Test and the NMC-Test affect the discrimination power of the test? and (4) What test-taking strategies do the first-year undergraduate students in the Faculty of Science at Mahidol University use while taking the C-Test, the MC-Test, the NMC-Test, and the NC-Test?

4.1 Results of the Original MC-Test and the NMC-Test

The original Modified C-Test (the MC-Test) is a test in which the *first half* or *the first part* of every *second word* is deleted and the students' task is to restore the deleted parts. When compared to the New Modified C-Test (NMC-Test) in this study, the *first half* or the *first part* of every *third word* is deleted in order to make this type of the language test be more suitable for Thai university students within an EFL context. So the present study was intended to investigate whether the original MC-Test yielded any different results from the NMC-Test with third-word deletion in measuring the English language proficiency of the first-year science students at Mahdiol University. The scores on the original MC-Test and the NMC-Test were then calculated by using the independent sample t-test, as can be seen in Table 7. The results of the present study reveal that there are statistically significant differences between the mean scores on the original MC-Test (M_{mc} = 62.21, SD = 14.29) and the mean scores on the NMC-Test taken by high-language-ability students, at the .05 level of significance, (M_{nmc} = 72.57, SD = 9.70), t (26) = -2.243, p < .017. For the low-language-ability group, the table below shows that there were no statistically significant differences at the .05 level of significance (p < .05)

Copyright by Mahidol University

between the mean scores of the original MC-Test (M_{mc} = 34.42, SD = 4.89) and the mean scores of the NMC-Test (M_{nmc} = 39, SD = 7.50), t (26) = -1.909. Although both high and low groups made higher scores on the NMC-Test than the original MC-Test, only the high-language-ability group showed statistically significant differences between these two types of the cloze test at the .05 level of significance ($p < .05$). This indicates that language teachers might use the original MC-Test and the NMC-Test as alternatives to measure the English language proficiency of first-year undergraduate science students within an EFL context. Further details of the differences in using these two types of cloze test will be discussed in the next chapter

Table 7: Comparison of the mean scores on the original MC-Test and the NMC-Test within the high-language-ability and low-language-ability groups

Group	Test Type	N	*M*	SD	t-value	*df*	Sig. (2-tailed)
High	The MC-Test	14	62.21	14.29	-2.243	26	.017*
	The NMC-Test	14	72.57	9.70			
Low	The MC-Test	14	34.42	4.89	1.909	26	.966
	The NMC-Test	14	39.00	7.50			

*$p < .05$

4.2 Results of the Original C-Test and the NC-Test

The C-Test is a test in which the *second half* or the *second part* of every *second word* is deleted, and the test-takers are required to restore all the deleted parts. Imitating the construction of the C-Test, the New C-Test (the NC-Test) was constructed in this study by deleting the *second half* or the *second part* of every *third word* to provide more clues for Thai university students studying English as foreign language. In this study, these two types of cloze test were further examined whether the original C-Test yielded any different results from the NC-Test with third-word deletion in measuring the English language proficiency of the first-year undergraduate students studying in the Faculty of Science at Mahidol University. So the scores on these two language tests were calculated by using the independent sample T-Test, as shown in Table 8. The findings of the present

Copyright by Mahidol University

study indicate that there were no statistically significant differences between the mean scores of the original C-Test (M_c = 83.35, SD = 9.85) and the mean scores of the NC-Test taken by the high-language-ability students at the .05 level of significance, (M_{nc} =77.53, SD =10.42), t (25) = -1.618, p > .05. The results of the low-language-ability group also showed that there were no statistically significant differences at the .05 level of significance between the mean scores on the original C-Test (M_c = 42, SD = 6.43) and the mean scores on the NC-Test at (M_{nc} = 44.38, SD =5.60), t (25) = -1.023, p > .05. Therefore, the C-Test is most likely to be a substitute for the NC-Test measuring the English language proficiency of the target students due to the fact that there were no statistically significant differences between the mean scores on the original C-Test and the NC-Test for both the high and low groups at the .05 level of significance (p >.05). Nonetheless, the differences between the original C-Test and the NC-Test with third-word deletion may result from the other factors: the type of deleted words and the deletion technique. Further information will be discussed in the next chapter.

Table 8: Comparison of the mean scores on the original C-Test and the NC-Test within the high-language-ability and low-language-ability groups

Group	Test Type	N	*M*	SD	t-value	*df*	Sig. (2-tailed)
High	**The C-Test**	14	83.85	9.85	-1.618	25	.059
	The NC-Test	13	77.53	10.42			
Low	**The C-Test**	14	42.00	6.43	-1.023	25	.158
	The NC-Test	13	44.38	5.60			

Copyright by Mahidol University

4.3 Results of Using Third-Word Deletion in the NMC-Test and the NC-Test

In this study, third-word deletion was employed on the new constructed cloze tests (the NC-Test and the NMC-Test) in order to examine what the effects of using third-word deletion in these two types of the cloze tests were. The scores on each type of cloze test were then calculated by using the independent sample t-test to see whether there were any statistically significant differences in using each type of cloze test in measuring the English language proficiency of the high-language-ability and low-language-ability groups. The findings of the study (see Table 9) reveal that there were statistically significant differences between the high and low groups on the mean scores of the four types of cloze tests at the .05 level of significance, ($p < .05$). This indicates that the four types of cloze tests, including the original C-Test, the NC-Test, the original MC-Test, and the NMC-Test, used in the study, could differentiate among first-year undergraduate science students at different language-ability levels.

Table 9: Comparison of the mean scores of the four cloze tests between the high-language-ability and low-language-ability groups

Test Type	Group	N	*M*	SD	t-value	*df*	Sig. (2-tailed)
C-Test	**High**	14	83.85	9.85	13.303	26	.000*
	Low	14	42.00	6.43			
NC-Test	**High**	13	77.53	10.42	10.096	24	.000*
	Low	13	44.38	5.60			
MC-Test	**High**	14	62.21	14.29	6.881	26	.000*
	Low	14	34.42	4.89			
NMC-Test	**High**	14	72.57	9.70	10.242	24	.000*
	Low	14	39.00	7.50			

***p* < .05**

Copyright by Mahidol University

Item analysis was also utilized in the present study to see whether using third-word deletion in the NC-Test and the NMC-Test had an influence on test discrimination power. Item analysis consisted of two calculations: "*Item Facility (IF)*" showing the difficulty level of the deleted items in these four cloze tests and "*Item Discrimination (ID)*" indicating how each cloze item differentiate the better students from the lower ones. In addition, the acceptable facility values should range from .20 to .80 and the discrimination indices should be equal to 0.30 or greater (Nuttal & Skurnik, 1969; Wongsothorn, 1996). However, discrimination indices higher than 0.25 were acceptable in the present study due to the fact that some deleted items could be readjusted (see Appendix H). The summary of the discrimination indices of the four cloze tests, as shown in Table 10, show that the more of the original C-Test items had better discrimination than those of the NC-Test. When comparing the original MC-Test with the NMC-Test, the total number of the NMC-Test deleted items with low discrimination power was slightly greater than that of the original MC-Test in assessing the general language skills of English.

Table 10: Item discrimination of the four types of cloze tests

Discrimination Index	C-Test	NC-Test	MC-Test	NMC-Test
≥ .76	12	9	5	3
.51 - .75	27	23	13	17
.26 - .50	35	24	39	39
≤ .25	26	44	43	41
Total	100	100	100	100

For the original C-Test, 26 deleted items (26%) out of the total number were considered to have low discrimination power. Most of the deleted items in the C-Test were reported to be too easy for these undergraduate science students due to the fact that this type of the language test possibly provides more chance of guessing. Only Item 82, as shown in Table 11, was too difficult for the first-year science students. Although the results of the present study indicate that the original C-Test could differentiate among these target students with high and low language abilities, some C-Test items should be

Copyright by Mahidol University

readjusted in order to make it more suitable for measuring English language proficiency within an EFL context.

Table 11: The C-Test items with low discrimination power

Item	Words	IF Total	ID	Item	Words	IF Total	ID
1	living	1.00	0.00	**51**	you	1.00	0.00
3	both	0.92	0.15	**53**	you	1.00	0.00
4	and	0.82	0.21	**55**	you	1.00	0.00
12	can	0.85	0.29	**63**	is	0.96	-0.08
25	that	0.50	0.15	**64**	very	0.96	0.08
26	have	0.96	0.08	**77**	very	0.96	0.08
29	that	0.82	0.07	**79**	years	0.92	0.15
30	other	1.00	0.00	**82**	stems	0.03[a]	0.07
33	and	0.89	0.57	**85**	is	0.64	0.14
35	thoughts	0.50	0.15	**90**	a	0.82	0.21
37	human	0.96	0.08	**92**	jogging	0.96	0.08
41	can	0.85	0.14	**97**	other	0.89	-0.07
50	they	0.92	0.15	**98**	running	0.92	0.15

[a]Item facility should range from 0.20-0.80, and item discrimination should be equal to 0.30 or greater.

Regarding the NC-Test with third-word deletion, there were 44 deleted items (44%) out of the total number that could not distinguish the high-language-ability students from the low-language-ability students. These NC-Test items with low discrimination power were reported to be too easy because the second half or the second part deletion may provide more chance of guessing. Only three deleted items (Items 3, 14, and 30) seemed to be too difficult for the first-year undergraduate science students (see Table 14). So there were 39 deleted items (39%) that were considered to be less suitable in measuring the English language proficiency of the target students. In addition, the results of the item analysis reported that using third-word deletion gave the NC-Test items lower discrimination power than the original C-Test in measuring the general language skills of English. Therefore, the different rate of deletion affected the discrimination power of the NC-Test in this present study.

Copyright by Mahidol University

Table 12: The NC-Test items with low discrimination power

Item	Words	IF Total	ID	Item	Words	IF Total	ID
3	provide	0.00 [a]	0.00	**50**	these	0.34	0.08
4	for	0.88	0.08	**51**	do	1.00	0.00
8	can	0.96	0.08	**52**	you	0.96	0.08
14	set	0.03 [a]	0.07	**53**	and	0.92	0.00
16	are	0.73	0.06	**63**	as	0.76	0.15
18	to	0.92	0.08	**64**	and	0.84	0.16
20	in	0.65	0.23	**65**	people	0.96	0.08
22	such	0.73	0.08	**71**	they	0.80	0.16
24	a	0.92	0.14	**72**	is	0.73	0.23
25	something	0.80	0.08	**73**	a	1.00	0.00
29	the	0.57	0.08	**74**	many	0.88	0.24
30	speech	0.00 [a]	0.00	**75**	to	1.00	0.00
32	ideas	0.38	0.16	**78**	the	0.88	0.08
33	Human	0.96	0.08	**79**	jogging	1.00	0.00
39	carefully	0.73	0.23	**85**	a	1.00	0.00
42	look	0.46	0.00	**90**	sports	0.88	0.24
45	They	0.92	0.16	**91**	biking	0.80	0.16
46	and	0.80	0.24	**93**	an	0.96	0.08
49	machines	0.61	0.16	**94**	such	0.88	0.08

[a]Item facility should range from 0.20-0.80, and item discrimination should be equal to 0.30 or greater.

In the original MC-Test, there were 43 deleted items (43%) out of the total number, which were considered to have low discrimination power as can be seen in Table 13. The 29 deleted items (29%) in the MC-Test were considered to be less suitable for measuring the English language proficiency of the target students because they were too easy. Only 14 deleted items (14%) (Items 9, 13, 14, 15, 21, 23, 46, 68, 80, 82, 83, 88, and 99) were too difficult for these first-year undergraduate science students. Consequently, the findings in Table 13 showed that some MC-Test items could not measure the language proficiency of English so they should be readjusted to make this type of language test appropriate for measuring the English language proficiency of EFL tertiary students.

Copyright by Mahidol University

Table 13: The MC-Test items with low discrimination power

Item	Words	IF Total	ID	Item	Words	IF Total	ID
2	both	0.32	0.07	**62**	world	0.60	0.07
3	and	0.89	0.22	**63**	is	1.00	0.00
9	tiny	0.10 [a]	0.21	**64**	very	0.78	0.00
10	and	0.64	0.57	**65**	smoking	0.67	-0.07
13	infected	0.14 [a]	0.28	**66**	many	0.75	0.07
14	even	0.10 [a]	0.21	**68**	serious	0.07 [a]	0.14
15	viruses	0.03 [a]	0.07	**71**	and	0.82	0.21
17	of	0.96	0.08	**77**	very	0.75	0.07
19	has	0.21	0.14	**80**	popularity	0.10 [a]	0.21
21	set	0.00 [a]	0.00	**81**	jogging	0.57	0.14
23	signs	0.03 [a]	0.07	**82**	stems	0.03 [a]	0.07
30	they	0.92	0.15	**83**	several	0.14 [a]	0.21
32	of	0.92	0.00	**85**	is	0.96	0.08
36	and	0.92	0.15	**86**	of	0.92	0.15
44	study	0.39	0.07	**88**	forms	0.10 [a]	0.21
46	carefully	0.07 [a]	-0.14	**89**	exercise	0.28	0.21
47	try	0.25	0.21	**90**	a	0.89	0.07
50	they	0.89	0.07	**91**	a	0.82	0.07
51	you	1.00	0.00	**93**	jogging	0.92	0.15
53	you	1.00	0.00	**96**	in	0.50	0.15
55	you	1.00	0.08	**99**	biking	0.03 [a]	0.07
59	smoking	0.64	0.00				

[a]Item facility should range from 0.20-0.80, and Item discrimination should be equal to 0.30 or greater.

For the NMC-Test with third-word deletion, the discrimination indices of the NMC-Test were slightly higher than the original MC-Test. The results of the present study also indicate that there were 41 deleted items (41%) in the NMC-Test that had low discrimination power. Thirty-two deleted items (32%) out of the total number were considered to be less suitable for measuring the English language proficiency because they were too easy. Only nine deleted items (Items 3, 14, 21, 23, 91, 92, 95, 99, and 100) were too difficult for the target students because the first-half or the first-part deletion possibly affected the difficulty of the deleted items. These target students may restore

Copyright by Mahidol University

some deleted items in the NMC-Test by using only grammatical knowledge. These limitations on using the NMC-Test in assessing overall language skills will be discussed in the next chapter

Table 14: The NMC-Test items with low discrimination power

Item	Words	IF Total	ID	Item	Words	IF Total	ID
2	and	0.92	0.15	**64**	and	0.89	0.22
3	provide	0.14 [a]	0.28	**65**	people	0.85	0.14
5	often	0.21	0.07	**71**	they	0.90	0.15
14	set	0.03 [a]	0.07	**72**	is	0.96	0.08
16	are	0.71	0.28	**73**	a	0.89	0.22
20	is	0.92	0.15	**74**	in	0.89	0.07
21	patient	0.00 [a]	0.00	**77**	many	0.85	0.1
23	such	0.00 [a]	0.00	**80**	from	0.67	0.21
24	bleeding	0.89	0.22	**85**	a	0.21	0.28
26	more	0.78	0.14	**87**	up	0.85	0.00
27	than	0.78	0.14	**89**	in	0.85	0.00
40	to	0.92	0.15	**91**	biking	0.07 [a]	0.14
45	they	0.92	0.15	**92**	quick	0.14 [a]	0.14
46	and	0.85	0.14	**93**	an	0.96	0.08
47	help	0.85	0.14	**95**	uses	0.03 [a]	0.07
51	do	0.96	0.08	**97**	In	0.78 [a]	0.00
52	you	1.00	0.00	**99**	rate	0.14	0.28
53	and	0.89	0.22	**100**	strengthens	0.03 [a]	0.07
63	as	0.96	0.08				

[a]Item facility should range from 0.20-0.80, and item discrimination should be equal to 0.30 or greater.

The results of the study then agree that using third-word deletion in the NC-Test and the NMC-test has an effect on the discrimination power of the test: high or low discrimination indexes. For example, the discrimination power of the NC-Test with third-word deletion was less than that of the original C-Test, whereas the discrimination power of the NMC-Test was slightly larger than the original MC-Test in measuring the English language proficiency of the target students. This occurrence seemed to be largely related to the word-deletion technique and the type of the deleted words (content words and functional words).

Copyright by Mahidol University

4.4 Results of Cloze Test-Taking Strategies Used by the First-Year Science Students

Test-taking strategies in the present study refers to the processes that the target students with high and low language abilities used while filling in the four types of the cloze test: the original C-Test, the NC-Test, the original MC-Test and the NMC-Test. The volunteer students' responses were then analyzed and categorized based on the latest categorization framework of cloze test-taking strategies by Sasaki (2000) in order to validate these four cloze tests as alternatives in measuring the English language proficiency within an EFL context, and to explore how these first-year science students responded to each type of cloze test. The results of the cloze test-taking strategies indicated by two raters are presented in Table 15. The findings below also reveal that both high and low-language-ability students used *Within Clause* strategy more frequently than other strategies. The students with high language ability could answer more items and used these four categories of test-taking strategies: *Within Clause*; *Across Sentences, Within Paragraph*; *Guessing*; and *Missing* more frequently than the low-level students. Moreover, the subjects in the low group often used only these three test-taking strategies: *Within Clause*; *Across Clause, Within Sentence;* and *Missing* while answering the four types of the cloze tests.

Copyright by Mahidol University

Table 15: The cloze test-taking strategies used by the high-language-ability and low-language-ability groups

Strategies	High-language-ability students							
	C-Test		NC-Test		MC-Test		NMC-Test	
	Rater 1	Rater 2	Rater 1	Rater 2	Rater 1	Rater 2	Rater 1	Rater 2
1. Within Clause	✓	✓	✓	✓	✓	✓	✓	✓
2. Across Clause, Within Sentence	✓	✓	✓	✓	✓	–	✓	✓
3. Across Sentences, Within Paragraph	✓	✓	✓	✓	✓	✓	✓	✓
4. Across Paragraphs, Within Text	–	–	–	–	✓	✓	✓	✓
5. Extratextual	✓	✓	–	–	✓	✓	✓	–
6. Guessing	✓	✓	✓	✓	✓	✓	✓	✓
7. Missing	✓	✓	✓	✓	✓	✓	✓	✓
8. Others	–	–	–	–	–	–	–	–
Strategies	**Low-language-ability students**							
	C-Test		NC-Test		MC-Test		NMC-Test	
	Rater 1	Rater 2	Rater 1	Rater 2	Rater 1	Rater 2	Rater 1	Rater 2
1. Within Clause	✓	✓	✓	✓	✓	✓	✓	✓
2. Across Clause, Within Sentence	✓	✓	✓	✓	–	–	✓	✓
3. Across Sentences, Within Paragraph	✓	✓	–	–	✓	✓	✓	–
4. Across Paragraphs, Within Text	–	–	–	–	✓	✓	✓	✓
5. Extratextual	–	–	–	–	✓	✓	–	–
6. Guessing	✓	✓	✓	✓	–	–	–	–
7. Missing	–	–	✓	✓	✓	✓	✓	✓
8. Others	–	–	–	–	–	–	–	–

Copyright by Mahidol University

The results calculated by Pearson's product moment, as shown in Table 16**,** show that there is a high correlation between the two inter-raters of the high and low-language-ability students' responses about test-taking strategies for each type of cloze test. For the C-Test and the NC-Test, there was a perfect correlation of the inter-rating results of the students' test-taking strategies (r = 1.000) in both the high- and low-language-ability groups. In the MC-Test, rater one indicated that the high-language-ability student taking the MC-Test used more strategies (Across Clause, Within Sentence) than rater two; however, the inter-rating result was highly correlated (r = .968). The inter-rating results also indicated that there was a moderate correlation of high student's responses (r = .761) and a high correlation of low student's responses (r = .911) on the NMC-Test. Rater one reported that the students taking the NMC-Test in both groups used more strategies than rater two.

Table 16: The correlation of inter-ratings of cloze test-taking strategies in the high-language-ability and low-language-ability groups indicated by two inter-raters

Test Type	High-language-ability		Low-language-ability	
	Rater 1	Rater2	Rater 1	Rater2
The C-Test	1.000	1.000	1.000	1.000
The NC-Test	1.000	1.000	1.000	1.000
The MC-Test	0.968	0.968	1.000	1.000
The NMC-Test	0.761	0.761	0.911	0.911

The test-taking strategies for the C-Test used by high- and low-language-ability students are presented in Table 17. Both high- and low-language-ability students said that they frequently used *Within Clause*; *Across Clause, Within Sentence*; *Across Sentences*, *Within Paragraph*; and *Guessing* while doing the C-Test. Only the student from the high-language-ability group used more *Extratextual* and *Missing* strategies than the student from the low-language-ability group. These two students with high and low-language-abilities also reported that taking the C-Test seemed like doing a language game, such as a crossword puzzle or hangman. They said, "I had a lot of fun playing such a puzzle game"

Copyright by Mahidol University

and "It seemed like playing hangman". However, neither used *Across Paragraphs, Within Text* strategies. Hence, the C-Test possibly measured language ability beyond the sentence level rather than overall language proficiency.

Table 17: Test-taking strategies for the C-Test transcribed as recorded

Strategies	The C-Test taken by high and low-language-ability students
1. Within Clause	**High:** "I had to find part of speech and the meaning of each item." **Low:** "I mean I had to find part of speech and the meaning of each item. For example, item 15, I could finally restore '**vir u s e s**' but I shouldn't have misspelled it."
2. Across Clause, Within Sentence	**High:** "For the third passage, the content is smoking. I used the contextual clue and the grammatical structure to check whether this sentence referred to smoking [pointing to item 58] or smoker [pointing to item 59]." **Low:** "I read for the main idea, and I also used contextual clues in each sentence."
3. Across Sentences, Within Paragraph	**High:** "I translated a few sentences at the beginning and at the end of the paragraph into Thai to find the main idea." **Low:** "I would read a few sentences at the beginning of the paragraph to find the main idea of the passage. In this way, I would understand what the passage was about."
4. Across Paragraphs, Within Text	**High:** ---- **Low:** ----
5. Extratextual	**High:** "The contents of some passages were so familiar to me that I could answer those items." **Low:** ----
6. Guessing	**High:** "Well, if the first part of each item is deleted, it is easy to guess. I had a lot of fun playing such a Puzzle game." **Low:** "I guessed if I didn't know the answer."
7. Missing	**High:** "However, if I didn't know the answers, I would leave them blank." **Low:** ----

The volunteer subjects with high-and low-language-ability used four test-taking strategies for the NC-Test (*Within Clause*; *Across Clause, Within Sentence*; *Guessing*; and *Missing*) more frequently than the other strategies. Only the high group reported the use of the *Across Sentences, Within Paragraph* strategy by reading a few sentences at the beginning of the paragraph to see what the passage was about (see Table 18). In addition, the students from the high- and low-language-ability groups said that they guessed the

Copyright by Mahidol University

answer by looking at the first part of the deleted words and counting the total number of letters. So the NC-Test in this study seemed to measure the language ability in the same way as the original C-Test, rather than the language proficiency of English.

Table 18: Test-taking strategies for the NC-Test transcribed as recorded

Strategies	The NC-Test taken by high and low-language-ability students
1. Within Clause	**High:** "Mostly, I focused on the meanings of the deleted words." **Low:** "Well, I used the clues preceding each item and also figured out what part of speech of each item should be."
2. Across Clause, Within Sentence	**High:** "I often used grammatical structures and part of speech, and translated each sentence with the missing words." **Low:** "I also used the context clues in the sentence."
3. Across Sentences, Within Paragraph	**High:** "I read only two or three sentences at the beginning of the paragraph, trying to find the main idea." **Low:** ----
4. Across Paragraphs, Within Text	**High:** ---- **Low:** ----
5. Extratextual	**High:** ---- **Low:** ----
6. Guessing	**High:** "I counted the number of the deleted letters on each item. Then, I guessed." **Low:** "First, I counted the number of the deleted letters of each item to help me guessed."
7. Missing	**High:** " But if I couldn't restore any items, I would ignore it." **Low:** "If I could not figure it out, I left it blank."

In the original MC-Test, the high and low-language-ability students mostly used the four strategies as can be seen in Table 19: *Within Clause*; *Across Sentences, Within Paragraph*; *Across Paragraphs, Within Text*; and *Missing*. Only the high-language-ability student reported using *Across Clause, Within Sentence* and *Guessing* strategies while taking the original MC-Test. However, the students with different language abilities also used more strategies by reading the whole passage in order to restore the deleted parts in this type of the cloze test. For example, "I read every sentence, line by line to see what the passage was about before filling in this language test". The findings of the study then show that the original MC-Test might be able to measure language ability beyond the sentence level.

Copyright by Mahidol University

Table 19:Test-taking strategies for the MC-Test transcribed as recorded

Strategies	The MC-Test taken by high and low-language-ability students
1. Within Clause	**High:** "But some items, such as this one [pointing to item 17] with one blank space and one letter, I knew immediately that it must be '**o f**'. For pronouns and prepositions, I could also restore immediately." **Low:** "I used the clues preceding the deleted words. I saw the word *example* so I knew that the word '**f o r**' could collocate with *example*. "
2. Across Clause, Within Sentence	**High:** "But sometimes, I had to analyze the deleted words to find which type of word is missing within each sentence." **Low:** **----**
3. Across Sentences, Within Paragraph	**High:** "…or read some parts of the test." **Low:** "I was looking the first two or three sentences of each passage."
4. Across Paragraphs, Within Text	**High:** "I read every sentence from the beginning to the end of the stories before filling in each blank." **Low:** "It was quite difficult. I read every sentence in the paragraph to see what it was about. For example, the first passage was so difficult that I could answer only a few items."
5. Extratextual	**High:** "I read most of the sentences in part because the test contents seemed to be general knowledge." **Low:** "I understood that it was about smoking so I could use my previous knowledge of vocabulary to help me fill in the deleted part correctly."
6. Guessing	**High:** "In addition, some answers I guessed were wrong." **Low:** **----**
7. Missing	**High:** "I couldn't restore about 15 to 16 items because I couldn't figure out." **Low:** "I could not guess any more so I left it blank."

The volunteer students from high- and low-language-ability groups reported using four test-taking strategies (*Within Clause; Across Clause, Within Sentence*; *Across Sentences, Within Paragraph*; and *Across Paragraphs, Within Text*) more frequently than the other strategies. Both students said that they read the story from the beginning to the end in order to fill in what was missing in the passage. Only the high-language-ability student indicated using *Extratextual* and *Guessing* strategies while responding to the NMC-Test. Therefore, the present study has found that the NMC-Test seems to measure language ability beyond the sentence level.

Copyright by Mahidol University

Table 20: Test-taking strategies for the NMC-Test transcribed as recorded

Strategies	The NMC-Test taken by high- and low-language-ability group
1. Within Clause	**High:** "I had to use the contextual clues to find the meanings of these words." **Low:** "I looked at the deleted word first. And then, I could think of the word to fill in the blank, such as in item 38."
2. Across Clause, Within Sentence	**High:** "I also analyzed part of speech of each keyword in each sentence." **Low: "**I had to use the contextual clues as well as grammatical structure of each sentence to help me think of a suitable word for each item."
3. Across Sentences, Within Paragraph	**High:** "I read most of the sentences in part." **Low:** "Just the same in that I had to read a few sentences at the beginning and at the end of the passage."
4. Across Paragraphs, Within Text	**High:** "I read the whole passage before restoring the deleted part." **Low:** "If the passages were difficult, such as passage one and passage two, I had to read through the passages once. And then, I tried to think of a suitable word for each item."
5. Extratextual	**High:** "I had some prior knowledge about some passages in the test so I felt familiar to the content of the test." **Low: ----**
6. Guessing	**High:** "I'm afraid that my answer might not be correct so I jotted down the word I guess in pencil for double-checking later." **Low: ----**
7. Missing	**High:** "Anyway, if I didn't know the answer for any item, I would leave it blank." **Low:** "Just left it blank. I didn't restore at all because I didn't know what the answer was."

To sum up, the four types of cloze test including the original C-Test, the NC-Test, the original MC-Test, and the NMC-Test, may be alternatives in measuring the English language proficiency of first-year undergraduate science students within an EFL context. Language teachers should take the text difficulty, the types of the deleted words, and the techniques of deletion into considerations due to the fact that these factors have an influence on the discrimination power of the test. Further details about using these four types of cloze test for assessing language proficiency will be discussed in the next chapter.

Copyright by Mahidol University

CHAPTER V
DISCUSSION

This chapter discusses and interprets the findings of the present study based on the four research questions: (1) Does the NMC-Test yield different results from the original MC-Test in measuring students' language proficiency? (2) Does the NC-Test yield different results from the original C-Test in measuring students' language proficiency? (3) Does using every third-word deletion of the NC-Test and the NMC-Test affect the discrimination power of the test? and (4) What test-taking strategies do the first-year undergraduate students in the Faculty of Science at Mahidol University use while taking the C-Test, the NC-Test, the MC-Test, and the NMC-Test?. Applications for the four cloze tests used in this study are presented.

5.1 Comparison of the Original MC-Test and the NMC-Test

The main purpose of this study was to compare the original MC-Test to the NMC-Test in measuring the English language proficiency of the first-year science students at Mahidol University. These science students with high- and low-language-ability ($N_1 = 56$) were asked to fill in all the deleted parts on these two types of the cloze test. In the high-language-ability group, the results show statistically significant differences at the .05 level of significance ($p < .05$) between the mean scores on the original MC-Test and those on the NMC-Test, while there were no statistically significant differences at the .05 level of significance ($p < .05$) between the mean scores on these two types of cloze test within the low-language-ability group. In addition, the findings reveal that these science students taking the NMC-Test ($M_{high} = 72.57$, $M_{low} = 39.00$) made higher scores than those taking the original MC-Test ($M_{high} = 62.21$, $M_{low} = 34.42$). These results are in agreement with the study of Wonghiransombat (1998), indicating that the 84 Thai postgraduate students got greater scores on the NMC-Test than on the original MC-Test. One possible explanation is that the NMC-Test with third-word deletion provides more clues for Thai undergraduate students studying English as a foreign language, so the NMC-Test seems to

Copyright by Mahidol University

be easier than the original MC-Test. Consequently, this study supports Wonghiransombat's findings that third-word deletion makes the NMC-Test more appropriate for Thai students for whom English is not the first language. Nonetheless, there are some factors, such as the type of deleted words and the deletion techniques that have an influence on the difficulty of the language test (Thongsa-nga, 1998; Wolter, 2002; Wonghiransombat, 1998). For example, some functional deleted items can be restored by using only grammatical or linguistic competence (Yamashita, 2003). Additionally, the first-half deletion in the original MC-Test makes the students use more strategies to fill in the deleted parts (Boonsathorn, 1987).

Regarding the reliability of the original MC-Test and the NMC-Test calculated by the Kuder-Richardson 21 formula, this study revealed that the reliability of the NMC-Test (r_{nmc} = .9417) was slightly higher than the original MC-Test (r_{mc} = .9285). These findings are consistent with Wonghiransombat's results which showed that the NMC-Test had greater reliability than the original MC-Test. Moreover, the present study also focuses on the content validity of these two language tests, which were constructed from the same text. The content of the four selected passages was verified in three ways: by the publishers, the subject teachers, and Fry's readability graph (1977, 2002) to confirm that these selected passages were suitable for the target science students. Wonghiransombat (1998) also recommended using various readability formulas as criteria to estimate the difficulty of passages extracted from newspapers. Language teachers also use the checklist to see whether the test content responds to the test purpose (Brown, 1996; Hughes, 2003). Nonetheless, the present study supports some other findings (Boonsathorn, 1987; Köberl & Sigott, 1996; Wonghiransombat, 1998), indicating that language teachers may use the original MC-Test and the NMC-Test as an alternative for assessing the English language proficiency of EFL university students due to the fact that these two language tests with the great reliability are easy to construct and to score.

Some studies (Prapphal, 1996; Sigott & Köberl, 1993) claimed that the original MC-Test seemed to measure language ability not beyond the sentence level. The students' responses in the present study show that the high- and low-language-ability students used strategies beyond the sentence level while taking the original MC-Test and the NMC-Test. For example, the volunteer students in the high-language-ability-group reported that

Copyright by Mahidol University

they used *Across Paragraphs, Within Text* strategy by reading line-by-line to understand the passage in both language tests. It is most likely that the MC-Test and the NMC-Test can be alternatives in measuring the English language proficiency of the target students. However, some test items in these two language tests with low discrimination power were too easy or too difficult because of the type of the deleted items: functional and content words. The functional words (such as an, in, and, their) are easier to restore than content words (such as communicate, provide, dangerous) (Abraham & Chapelle, 1992; Jafapur, 1999; Klein-Braley, 1997; Thongsa-nga, 1998; Wolter, 2002; Wonghiransombat, 1998). As a result, language teachers should take the above factors into considerations in order to make the original MC-Test and the NMC-Test more suitable for first-year undergraduate science students.

5.2 Comparison of the Original C-Test and the NC-Test

The present study was intended to examine whether the original C-Test yielded any different results from the NC-Test in measuring the English language proficiency of Thai first-year undergraduate students studying in the Faculty of Science at Mahidol University. These target students with high- and low-language-ability (N_2 = 54) were asked to fill in all the deleted parts in the original C-Test and the NC-Test. The results found no statistically significant differences at the .05 level of significance ($p < .05$) from the mean scores on the original C-Test and those on the NC-Test in both the high- and low-language-ability groups. It is most likely that the C-Test may be a substitute for the NC-Test in measuring English language proficiency. When comparing the mean scores on the target students within the same language ability groups, the high-language-ability students taking the original C-Test made slightly higher scores (M_c = 83.35) than those taking the NC-Test (M_{nc} = 77.53). In the low-language-ability group, the mean score on the NC-Test (M_{nc} = 44.38) was slightly higher than the mean score of the original C-Test (M_c = 42.00). Both the original C-Test and the NC-Test provide several choices for language teachers in assessing the language proficiency of EFL university students. Nonetheless, these findings of the present study are in agreement with other previous research (Cleary, 1988; Jafapur, 1999; Thongsa-nga, 1998; Wolter, 2002) reporting that some C-Test items, especially the functional words had low discrimination power.

Copyright by Mahidol University

Therefore, the type of deleted words (content words and functional words) has an effect on the test difficulty (Klein-Braley, 1997). Weir (1990) adds that the C-Test provides more chances of guessing due to the second-half or the second-part deletion.

The present study shows that the reliability of the original C-Test (r_c = .9648) was slightly greater than the reliability of the NC-Test with third-word deletion (r_{nc} = .9470) calculated by Kuder-Richardson 21 formula. However, Thongsa-nga (1998) found that the reliability of the NC-Test with third-word deletion was higher than the other forms of the NC-Tests. The content of the four selected passages were graded by the publishers, the subject teachers, and Fry's readability graph (1977, 2002) to ensure that these selected texts were suitable for these undergraduate science students of EFL. In addition, Thonsa-nga suggested using various passages from the *Student Weekly* newspaper, whose difficulty levels were estimated by Flesch-Kincaid and Fry's readability formulas in order to construct an English proficiency test for Mathayomsuksa Six students (equivalent to twelfth-grade students).

In addition, the results of the test-taking strategies for the C-Test and the NC-Test reveal that both high- and low-language ability students frequently used the *Across Sentences, Within Paragraph* strategies. These two types of cloze test thus seemed to be related to language ability at the syntactic level. As a consequence, these findings support some previous studies (Cohen, Segal, & Weiss, 1984; Jafapur, 1995; Sigott & Köberl, 1993) showing that the C-Test is considered to be less suitable for measuring English language proficiency of EFL university students. Conversely, many other studies (Babaii & Ansary, 2001; Cohen, Segal, & Weiss, 1984; Connelly, 1997; Dörnyei & Katona, 1992, 1993; Klein-Braley, 1985, 1997) still claimed that the C-Test was reliable in measuring the proficiency of second and foreign language learners, and the C-Test was easy to score and construct. To sum up, language teachers who want to make use of these two types of cloze test should take more consideration in word selection and deletion techniques that are suitable for the students' language ability.

Copyright by Mahidol University

5.3 Effects of Using Third-Word Deletion in the NC-Test and the NMC-Test

The present study was designed to find whether using third-word deletion in the NC-Test and the NMC-Test affected test discrimination power. The results found that there were statistically significant differences at the .05 level of significance ($p < .05$) between the high- and low-language-ability students taking each type of the four cloze tests. The four types of cloze test then could differentiate among the first-year undergraduate science students at different language-ability levels. Nonetheless, using third-word deletion in the NC-Test and the NMC-Test may result in high or low discrimination indices. Some test items with low discrimination power indicated that the test items were too easy or too difficult for the target students. When comparing the original C-Test with the NC-Test, the findings of the item analysis (see Appendix H) showed that only 26 C-Test items (26%) had low discrimination power. In the NC-Test, there were 39 deleted items (39%) which were considered to be less suitable for measuring English language proficiency due to the fact that these items were too easy for the target students. Weir (1990) added that the second half or the second part deletion in these two language tests provides more chances of guessing. The volunteer students with high- and low-language-ability also reported that they could fill in the answer by focusing on the clues in the first part of the deleted words. For instance, a student said, “I looked at the first part of the deleted words. It gave me a clue what was missing”

Moreover, Klein-Braley (1997) added, “the actual difficulty of the test constructed from the same text varied according to the proportion of function words deleted” (p. 59). The more content words each sentence of a passage contains, the more difficult the passage is for the fill-in activity. Many studies also agree that deleted functional words can be restored more easily than deleted content words (Cleary, 1988; Jafapur, 1999; Klein-Braley, 1997; Thongsa-nga, 1998; Weir, 1990; Wolter, 2002; Yamashita). In the original C-Test, there were 69 content words and 31 functional words, while the NC-Test consisted of 63 content words and 37 functional words, as can be seen in Table 21. So third-word deletion in the NC-Test, providing more clues for these EFL tertiary students, made the NC-Test easier than the original C-Test with second half or second part deletion.

Copyright by Mahidol University

Table 21: Number of content words and functional words in the four types of cloze test

Content Word	C-Test	NC-Test	MC-Test	NMC-Test
Noun	29	28	29	28
Verb	19	15	19	15
Adjective	13	11	13	11
Adverb	8	9	8	9
Total	**69**	**63**	**69**	**63**
Functional Word	**C-Test**	**NC-Test**	**MC-Test**	**NMC-Test**
Pronoun	8	10	8	10
Preposition	13	15	13	15
Conjunction	6	5	6	5
Article	3	7	3	7
Negative	1	-	1	-
Total	**31**	**37**	**31**	**37**

Regarding the effects of using third-word deletion in the NMC-Test, the results of the study show that the discrimination power of the NMC-Test was slightly greater than the original MC-Test in measuring the students' language proficiency. Wonghiransombat (1998) also contends that the discrimination indices of the NMC-Test in her study were higher than the original MC-Test. In the original MC-Test, there were 29 deleted items (29%) that seemed to be less suitable in measuring language proficiency in English within an EFL context because these items were too easy for these target students. Only 13 MC-Test items (13%) seemed to be too difficult. Compared to the NMC-Test with third-word deletion, the results of the item analysis showed that there were 41 deleted items with low discrimination power in the NMC-Test. However, 32 deleted items (32%) were too easy for the target university students, and nine NMC-Test items (9 %) were considered to be too difficult.

It is possible that the original MC-Test seemed to be more difficult than the NMC-Test with third-word deletion because there are more deleted content words available in the language tests. The data, as shown in Table 21, reveal that the original MC-Test is composed of the 69 content words and 31 functional words, whereas the NMC-Test with

Copyright by Mahidol University

third-word deletion contains 63 content words and 37 functional words. So third-word deletion in this study made the NMC-Test easier than the original MC-Test. Moreover, these findings show that different deletion techniques affect each type of cloze test used in this study, especially its discrimination power. So the results of the present support some research findings (Jafapur, 1995; Thongsa-nga, 1998), indicating that changing rates of deletion and/or the starting points has an influence on language tests. On the contrary, Wonghiransombat (1998) claimed that third-word deletion in the NMC-Test in her study did not affect the reliability, validity, and difficulty of the test.

In conclusion, the findings of the present study show that using third-word deletion in the NC-Test and the NMC-Test resulted in test discrimination power. Language teachers should thus be aware that using third-word deletion provides more clues for EFL tertiary students, since the third-word deletion probably makes the cloze tests more appropriate or lowers the discrimination power.

5.4 Cloze Test-Taking Strategies Used by the Volunteer Students

An investigation into test-taking strategies were also carried out in the present study to see how these first-year science students with high- and low-language-ability responded to the original C-Test, the NC-Test, the original MC-Test, and the NMC-Test. The results about test-taking strategies may provide more evidence for language teachers to validate these four types of cloze tests used in this study. The interview data obtained from the volunteer students (N_3 = 8) were analyzed and tabulated, based on the latest categorization framework of cloze test-taking strategies by Sasaki (2000). These findings reveal that the volunteer students reported using the *Within Clause* strategy most. So this study is in agreement with the investigation of Sasaki, indicating that the EFL Japanese students frequently used the *Within Clause* strategy to fill in familiar and unfamiliar cloze passages. Yamashita (2003) also found that the 12 EFL Japanese students often used the *Across Paragraphs, Within Text* strategy most while taking the rational cloze test. However, the high-language-ability students in the present study could answer more items correctly and used these four test-taking categories: *Within Clause*; *Across Sentences, Within Paragraph*; *Guessing*; and *Missing* more frequently than the students in the low-language-ability group. The low-language-ability students mostly used these three test-

Copyright by Mahidol University

taking strategies: *Within Clause*; *Across Clause, Within Sentence;* and *Missing* while answering the four types of cloze test.

In the C-Test, this present study reveals that these students frequently used *Within Clause*; *Across Clause, Within Sentence*; *Across Sentences*, *Within Paragraph*; and *Guessing* while taking the C-Test. Only a volunteer student of the high-language-ability group reported using the *Extratextual* and *Missing* strategies. It may be because their background knowledge and their experience in language learning are somewhat different. If the passage is familiar to the students' past experience, they could restore test items more easily than unfamiliar ones. Nonetheless, the students' responses showed that the C-Test in this study seemed to measure language ability at the syntactic level rather than the discourse level. These findings are consistent with many previous studies (Cohen, Segal, & Weiss, 1984; Jafapur, 1995, 1999; Prapphal, 1994; Sigott & Köberl, 1993), indicating that the measurement of the C-Test is related to lexical and grammatical competence in English, whereas some other studies (Babaii & Ansary, 2001; Cohen, Segal, & Weiss, 1984; Connelly, 1997; Dörnyei & Katona, 1992, 1993; Klein-Braley, 1985, 1997; Raatz & Klein-Braley, 1981) argue that the C-Test could assess the proficiency of the target language. Consequently, language teachers could use the C-Test for measuring specific knowledge, such as testing vocabulary and grammar.

These volunteer students frequently used four test-taking strategies on the NC-Test: *Within Clause*; *Across Clause, Within Sentence*; *Guessing*; and *Missing* more frequently that the other strategies. These test-taking strategies were not beyond the sentence level. Only a high-language-ability student reported the use of the *Across Sentences, Within Paragraph* strategy by reading a few sentences at the beginning of the paragraph to see what the passage was about. However, the findings in this study indicated that the NC-Test with third-word deletion and the C-Test could be used to measure EFL students in a similar way. This is because the NC-Test imitated the construction of the original C-Test. In addition, the results reveal that this type of language test provides more chances of guessing. For example, the volunteer students guessed the answer by focusing on the first-part of the deleted words or counting the total number of letters in the deleted words. So this study is consistent with Thongsa-nga's

Copyright by Mahidol University

results, indicating that the NC-Test with third-word deletion is suitable for measuring the English language proficiency of twelfth graders.

The findings in the present study found that the high- and low-language-ability students used *Within Clause*; *Across Clause, Within Paragraph*; *Across Paragraphs, Within Text*; and *Missing* most frequently while taking the original MC-Test. Only a high-language-ability student reported using the *Across Clause, Within Sentence* and *Guessing* strategies while answering the original MC-Test. However, the students with different language abilities also used more strategies by reading the whole passage in order to restore the deleted parts in this cloze test. For example, a volunteer student said, "I read every sentence, line-by-line, to see what the passage was about". Boonsathorn (1987) also argues that the ESL learners taking the original MC-Test require more strategies than those taking the original C-Test. So the results of the study agree that the MC-Test could measure the language proficiency of the target students even though this type of cloze test seems to be too difficult for non-native-speaking students (Sigott & Köberl, 1993).

The findings also reveal that the students with high- and low-language ability mostly used the following strategies: *Within Clause*; *Across Clause*, *Within Sentence*; *Across Sentences*, *Within Paragraph*; and *Across Paragraphs*, *Within Text* while taking the NMC-Test. It was found that the students read the story from the beginning to the end, which helped them understand the passages of the NMC-Test better. Nonetheless, these results confirm that the NMC-Test, imitating the construction of the original MC-Test, could measure language ability beyond the sentence level. The NMC-Test in this study can be one of several choices for language teachers to assess the proficiency of the EFL students; however, a teacher should take consideration the types of the deleted words and the deletion techniques.

In summary, these four types of language tests provide many alternative ways to assess English language proficiency of EFL tertiary students in the science program. Both the C-Test and the NC-Test in this study seem to be suitable for measuring specific language ability, such as grammar and vocabulary, due to the fact that the volunteer students used only the test-taking strategies at the sentence level, such as *Within Clause*; *Across Clause*, *Within Sentence*, and *Across Sentences*, *Within Paragraph*. Conversely, the MC-Test and the NMC-Test were considered as suitable to measure the English

Copyright by Mahidol University

language proficiency of non-native-speaking undergraduates in the science program because the volunteer students used test-taking strategies beyond the sentence level or at the discourse level, such as the *Across Paragraphs*, *Within Text* strategy. As a consequence, language teachers should be aware that different types of deleted words and different deletion techniques can have an influence on these four types of cloze tests.

Copyright by Mahidol University

CHAPTER VI
CONCLUSION AND RECOMMENDATIONS

Chapter six provides the conclusion of the present study, the recommendations for pedagogical applications, and the recommendations for further study.

6.1 Conclusion

The main purpose of the present study was to investigate the four types of cloze test by comparing the original C-Test with the New Constructed C-Test (the NC-Test, and the original Modified C-Test (the MC-Test) with the New Modified C-Test (the NMC-Test) in measuring the English language proficiency of first-year undergraduate students studying in the science program at Mahidol University. The 110 students in high- and low-language-ability groups were divided into four sub-groups taking four types of the cloze test. Moreover, this study aimed to examine the test-taking strategies these target students used to restore deleted items in the C-Test, the NC-Test, the MC-Test, and the NMC-Test in measuring English language proficiency. The results of the study may provide an alternative way for language teachers to assess the English language proficiency of non-native-speaking university students.

The statistical devices used in this study were the mean and standard deviation reporting the average scores for the four types of the cloze test in measuring the students' English language proficiency. The independent sample t-test was employed to analyze whether there were any significant differences between the students' test scores on the original C-Test and the NC-Test, and between the original MC-Test and the NMC-Test. In addition, item analysis was also used in this study to see whether the third-word deletion in the NC-Test and the NMC-Test affect the test discrimination power. For the test-taking strategies, the students' responses obtained from group interview were reported by calculating the frequencies of what test-taking strategies these volunteer students used while taking the four cloze tests.

Copyright by Mahidol University

The mean scores on the cloze tests were calculated by the independent sample t-test to find whether the original C-Test yielded any different results from the NC-Test in measuring the students' English language proficiency within high- and low-language-ability groups. These findings show that there are no statistically significant differences at the .05 level of significance ($p < .05$) between the mean scores on the C-Test and the NC-Test within the high- and low-language-ability groups. This indicates that the NC-Test can be a substitute for the original C-Test in both high- and low-language-ability groups. In addition, the test scores on the original MC-Test and the NMC-Test, calculated by the independent sample t-test, showed that there were statistically significant differences between the mean scores on the MC-Test and the NMC-Test in the high-language-ability group at the .05 level of significance ($p < .05$), whereas these findings found no statistically significant differences at $p < .05$ between the mean scores on these two language tests in the low-language-ability students. However, the high- and low-language-ability students made higher scores on the NMC-Test than on the original MC-Test.

In this study, third-word deletion was employed in the NC-Test and the NMC-Test to provide more clues for the EFL university test-takers. These findings reveal that using third-word deletion has an influence on the discrimination power of the test, such as high or low discrimination power. When comparing the original C-Test with the NC-Test, third-word deletion gave lower discrimination power to the NC-Test than to the original C-Test. The discrimination indices of the NMC-Test were slightly higher than those of the original MC-Test. Another factor that affected the four types of cloze test in this study was the type of deleted words. For example, deleted functional words could be restored by using only linguistic or grammatical competence. Consequently, language teachers should be aware that different types of deleted words and different deletion techniques can have an influence on these four types of cloze test in measuring the English language proficiency of EFL tertiary students.

The interview data in this study reveal that the original C-Test and the NC-Test can measure language ability at the sentence level, so these two language tests are suitable for measuring specific knowledge of English, such as grammar or vocabulary. The students' responses also show that these volunteer students used the *Across Sentences*,

Copyright by Mahidol University

Within Paragraph strategy, indicating that these two type of cloze test can measure the English language proficiency of the target students within an EFL context.

6.2 Recommendations for Further Study

This comparative study was limited to the first-year undergraduate science students studying in the Faculty of Science at Mahidol University. Further studies should be conducted to examine the four types of the cloze tests: the original C-Test, the NC-Test, the original MC-Test, and the NMC-Test taken by university students in different academic fields. Moreover, there should be a further study using the different types of the cloze test in different testing situations, such as reading proficiency. Boonsathorn (2000) also proposed the latest form of the cloze test, the S-Test (Semantic/Syntactic Test) by deleting half or part of every second content words. Example 22 provides the S1-Test in which the *first half* or the *first part* of every second *content word* is deleted, and the S2-Test in which the *second half* or the *second part* of every *content word* is deleted.

Example 22: S1-Test vs. S2-Test

S1-Test	**S2-Test**
Tom was a student who wasn't interested in studying. He preferred to _ _ **ve** a good _ _ **me**. Naturally, when he _ _ **ok** his examinations, he didn't _ _ **t** good _ _ _ **ks**. Since he knew his _ _ _ **her** would be _ _ _ **ry** with him, he sent a _ _ _ _ **gram** to his brother's _ _ _ **se**.	Tom was a student who wasn't interested in studying. He preferred to **ha** _ _ a good **ti** _ _. Naturally, when he **to** _ _ his examinations, he didn't **g** _ _ good **ma**_ _ _. Since he knew his fat _ _ _ would be **an** _ _ _ with him, he sent a **tele** _ _ _ _ to his brother's **ho** _ _ _.
(Extracted from Boonsathorn, 2000, p. 68)	(Extracted from Boonsathorn, 2000, p. 68)

In addition, an investigation of the test-taking strategies for cloze tests, including the S-Test, should be carried out by using a greater number of students to better establish the validity of the research findings. For English language teachers, these findings provide several choices in measuring English language proficiency. Language teachers may use the C-Test and the NC-Test in measuring the specific knowledge of English language

Copyright by Mahidol University

ability, such as vocabulary and grammar, and use the MC-Test and the NMC-Test as a part of testing English language proficiency of EFL university students.

Copyright by Mahidol University

REFERENCES

Abraham, R. G., & Chapelle, C. A. (1992). The meaning of cloze test scores: An item difficulty perspective. *The Modern Language Journal, 76*, 468-479.

Ackart, P. (1997). *Concepts and comments: A reader for students of English as a second language* (2nd ed.). Forth Worth: Harcourt Brace.

Aitken, K. G. (1977). Using cloze procedure as an overall language proficiency test. *TESOL Quarterly, 11*, 59-67.

Alderson, J. C. (1979). The cloze procedure and proficiency in English as a foreign language. *TESOL Quarterly, 13*, 219-227.

Alderson, J. C. (1980). Native and nonnative speaker performance on cloze tests. *Language Learning, 30*, 59-76.

Alderson, J. C. (1983). The cloze procedure and proficiency in English as a foreign language. In J. W. Oller, Jr., (Ed.), *Issues in language testing research* (pp. 205-217). Massachusetts: Newbury House.

Alderson, J. C. (2000). *Assessing reading*. Cambridge: Cambridge University Press.

Anderson, G. (1998). *Fundamentals of Educational Research* (2nd ed.). London: Falmer Press.

Anderson, J. (1976). *Psycholinguistics experiments in foreign language testing*. Queensland: University of Queensland Press.

Anderson, J. (1983). Lix and Rix: Variations on a little-known readability index. *Journal of Reading, 26*, 490-497.

Babaii, E., & Ansary, H. (2001). The C-Test: A valid operationalization of reduced redundancy principle? *System, 29*, 209-219.

Bachman, L. F. (1982). The trait structure of cloze test scores. *TESOL Quarterly, 16*, 61-70.

Bachman, L. F. (1985). Performance on cloze tests with fixed-ratio and rational deletions. *TESOL Quarterly, 16*, 61-70.

Copyright by Mahidol University

Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.

Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice*. Oxford: Oxford University Press.

Bailey, K. M. (1998). *Learning about language assessment: Dilemmas, decisions, and directions*. Pacific Grove: Heinle & Heinle.

Baldauf, R. B., Jr., & Propst, I. K., Jr. (1979). Matching and multiple-choice cloze tests. *Journal of Education Research, 72*, 321-326.

Boonsathorn, S. (1987). C-Tests, proficiency, and reading strategies in ESL. *Dissertation Abstracts International, 48*(05), 1111. (UMI No. 0560701)

Boonsathorn, S. (1990). C-Test/MC-Test: A new instrument for measuring English language proficiency. *PASAA, 20*, 46-50.

Boonsathorn, S. (2000). S-Tests: A rational C-Test for advanced students. *Thai Journal of Development Administration, 40*, 65-74.

Bradshaw, J. (1990). Test takers' reaction to a placement test. *Language Testing, 7*, 13-30.

Brown, D. (1983). Conversational cloze tests and conversational ability. *English Language Teaching Journal, 37*, 158-161.

Brown, J. D. (1980). Relative merits of four methods for scoring cloze tests. *The Modern Language Journal, 64*, 311-317.

Brown, J. D. (1983). A closer look at cloze tests with fixed-ratio and rational deletions. *TESOL Quarterly, 19*, 535-556.

Brown, J. D. (1996). *Testing in language programs*. New Jersey: Prentice-Hall.

Chapelle, C. A., & Abraham, R. G. (1990). Cloze method: What difference does it make? *Language Testing, 7*, 121-146.

Chavanachat, P. (1986). The use of a cloze test as a measure of English proficiency for Thai university graduates. *PASAA, 16*, 26-35.

Chàvez-Oller, M. A., Chihara, T., Weaver, K. A., & Oller, J. W., Jr. (1985). When are cloze items sensitive to constraints across sentences? *Language Learning, 35*, 181-206.

Cleary, C. (1988). The C-Test in English: Left-hand deletions. *RELC Journal, 19*(2), 26-38.

Copyright by Mahidol University

Cohen, A. D. (1980). *Testing language ability in the classroom*. Rowley: Newbury House.

Cohen, A. D. (1984). On taking language tests: what the students report. *Language Testing, 1*, 70-81.

Cohen, A. D. (1994). *Assessing language ability in the classroom* (2nd ed.). Boston: Heinle & Heinle.

Cohen, A. D. (1998). *Strategies in learning and using a second language*. London: Longman.

Cohen, A. D., Segal, M., & Weiss, R. (1984). The C-Test in Hebrew. *Language Testing, 1*, 221-225.

Cohen, L., Manion, L., & Morrison, K. (2000). *Research methods in education*. London: RoutledgeFalmer.

Connelly, M. (1997). Using C-Test in English with postgraduate students. *English for Specific Purposes, 16*, 139-150.

Crawley, J., & Mountain, L. (1995). *Strategies for guiding content reading* (2nd ed.). Bonton: Allyn and Bacon.

Dörnyei, Z., & Katona, L. (1992). Validation of the C-Test amongst Hungarian EFL learners. *Language Testing, 9*, 187-206.

Dornyei, Z., & Katona, L. (1993). The C-Test: A teacher friendly way to test English language proficiency. *The English Teaching Forum, 31*, 35-40. Retrieved January 1, 2002, from http://exchanges.state.gov/forum/vols/vol31/mo1/p35.htm

Duffelmeyer, F. A. (1983). The effect of grade level on cloze test scores. *Journal of Reading, 26*, 436-441.

Fotos, S. S. (1991). The cloze test as an integrative measure of ESL proficiency: A substitute for essays on college entrance examinations? *Language Learning, 41*, 331-336.

Fry, E. B. (1977). Fry's readability graph: Clarifications, validity, and extension to level 17. *Journal of Reading, 17*, 242-252.

Fry, E. B. (1989). Reading formulas maligned but valid. *Journal of Reading, 32*, 292-297.

Fry, E. B. (1990). A readability formula for short passages. *Journal of Reading, 33*, 594-597.

Fry, E. B. (2002). Readability versus leveling. *The Reading Teacher, 56*, 286-291.

Copyright by Mahidol University

Gregg, J. Y., & Russell, J. (1990). *Past, present and future: A reading-writing text* (3rd ed.). California: Wadsworth publication.

Grotjahn, R. (1986). Test validation and cognitive psychology: Some methodological considerations. *Language Testing, 3*, 159-185.

Heaton, J. W. (1975). *Writing English language tests*. London: Longman.

Hinofotis, F. B., & Snow, B. G. (1980). An alternative cloze testing procedure: Multiple-choice format. In J. W. Oller, Jr. (Ed.), *Research in language testing* (pp. 129-133). Massachusetts: Newbury House.

Hinofotis, F. B. (1980). Cloze as an alternative method of ESL placement and proficiency testing. In J. W. Oller, Jr. (Ed.), *Research in language testing* (pp. 121-128). Massachusettes: Newbury House.

Hughes, A. (2003). *Testing for language teachers* (2nd ed.). Cambridge: Cambridge University Press.

Jafapur, A. (1995). Is C-Test superior to cloze? *Language Testing, 12*, 194-216.

Jafapur, A. (1999). Can the C-Test be improved with classical item analysis? *System, 27*, 79-89.

Jitendra, A. K., & Rohena-Diaz, E. (1996). Language assessment of students who are linguistically diverse: Why a discrete approach is not the answer. *The School Psychology Review, 25*, 40-56.

Jonz, J. (1987). Textual cohesion and second-language comprehension. *Language Learning, 37*, 409-438.

Jonz, J. (1990). Another turn in the conservation of language complexity and cultural background of a text. *TESOL Quarterly, 24*, 61-83.

Kaczmarek, C. M. (1980). Scoring and rating essay tasks. In J. W. Oller, Jr. (Ed.), *Research in language testing* (pp. 151-159). Massachusetts: Newbury House.

Klein-Braley, C. (1985). A cloze up on the C-Test: A study in the construct validation of authentic tests. *Language Testing, 2*, 76-104.

Klein-Braley, C. (1997). C-Test in the context of reduced redundancy testing: An appraisal. *Language Testing, 14*, 47-84.

Klein-Braley, C., & Raatz, U. (1984). A survey of research on the C-Test. *Language Testing, 1*, 134-146.

Copyright by Mahidol University

Köberl, J., & Sigott, G. (1996). Deletion patterns in C-Test: Native speakers vs. foreign learners. *Language Testing Update, 20*, 45-48.

Lange, D. L., & Clausing, G. (1981). An examination of two methods of generating and scoring cloze tests with students of German on three levels. *The Modern Language Journal, 65*, 254-261.

Leu, D. J., & Kinzer, C. K. (1995). *Effective reading instruction, K-8* (3rd ed.). Englewood Cliffs: Merrill.

Lyman, B. H. (1998). *Test scores and what they mean* (6th ed.). Boston: Allyn and Bacon.

McKenna, M. C., & Layton, K. (1990). Concurrent validity of cloze as a measure of intersentential comprehension. *Journal of Educational Psychology, 82*, 372-377.

McNamara, T. (2000). *Language testing*. Oxford: Oxford University Press.

Markham, P. L. (1985). Rational deletion cloze and global comprehension in German. *Language Learning, 35*, 423-430.

Markham, P. L. (1987). Rational deletion cloze processing strategies: ESL and native English. *System, 15*, 303-311.

Mullen, K. A. (1979). An alternative to the cloze test. In C. A. Yorio, K. PerKins, & J. Schachter (Eds.), *On TESOL 79: The learner in focus selected papers from the 13th Annual Convention Boston, Mass. February 27 - March 4, 1979* (pp. 187-205). Washington, D. C.: TESOL.

Nuttal, D. L., & Skurnik, L. S. (1969). *Examination and item analysis manual*. England and Wales: National Foundation For Educational Research.

O' Malley, J. M., & Chamot, A. U. (1990). *Learning strategies in second language acquisition*. Cambridge: Cambridge University Press.

Oller, J. W., Jr. (1972). Scoring methods and difficulty levels for cloze tests of proficiency in ESL. *The Modern Language Journal, 56*, 151-158.

Oller, J. W., Jr. (1973). Cloze tests of language proficiency and what they measure. *Language Learning, 23*, 105-118.

Oller, J. W., Jr. (1979). *Language tests at school: A pragmatic approach*. London: Longman.

Oller, J. W., Jr., & Conrad, C. (1971). The cloze technique and ESL proficiency. *Language Learning, 21*, 183-195.

Copyright by Mahidol University

Oller, J. W., Jr., & Inal, N. (1971). A cloze test of English prepositions. *TESOL Quarterly, 5*, 315-326.

Porter, D. (1975). Modified cloze procedure: A more valid reading comprehension test. *English Language Teaching Journal, 30*, 151-155.

Prapphal, K. (1994). A study of the C-Test and the X-Test performed by first-year science-oriented university students. *PASAA, 24*, 16-23.

Prapphal, K. (1996). An investigation of the general and academic English X-Tests in measuring grammatical competence of Thai science students. *PASAA, 26*, 73-79.

Raatz, U., & Klein-Braley, C. (1981). The C-Test--a modification of the cloze procedure. In T. Culhane, C. Klein-Braley & D. K. Stevenson (Eds.), *Practice and problem in language testing* (pp. 113-138). Colchester: University of Essex.

Raatz, U., & Klein-Braley, C. (2003). *Introduction to language testing and C-Tests*. Retrieved February 20, 2003, from www.uni-duisburg.de/fb3/angling/forschung/howtodo.htm

Richards, J. C., Platt, J., & Platt, H. (1993). *Longman dictionary of language teaching and applied linguistics* (3rd ed.). Essex: Longman.

Saetung, S. (1984). *The application of Rix readability index to determine readability levels of supplementary English reading books for lower secondary school students.* Master's thesis, Kasetsart University, Bangkok, Thailand.

Sasaki, M. (2000). Effects of cultural schemata on student's test-taking processes for cloze tests: A multiple data source approach. *Language Testing, 17*, 85-114.

Sawatdirakpong, S. (1980). *The effectiveness of cloze in teaching reading comprehension.* Master's thesis, Mahidol University, Bangkok, Thailand.

Shanahan, T., Kamil, M. L., & Tobin, A. W. (1982). Cloze as a measure of intersentential comprehension. *Reading Research Quarterly, 17*, 229-255.

Sigott, G., & Köberl, J. (1993). Validating the X-Test. *Language Testing, 14*, 53-58.

Spargo, E. (2002a). *Timed reading plus: 25 two-part lessons with questions for building reading speed and comprehension book 4*. Singapore: Jamestown.

Spargo, E. (2002b). *Timed reading plus: 25 two-part lessons with questions for building reading speed and comprehension book 5*. Singapore: Jamestown.

Copyright by Mahidol University

Spolsky, B. (1973). What does it mean to know a language, or how do you get somebody to perform his competence? In J. W. Oller, Jr. & J. R. Richards (Eds.), *Focus on the learner* (pp. 164-176). Rowley: Newbury House.

Standal, T. C., & Betza, R. E. (1990). *Content area reading teachers, texts, students*. New Jersey: Prentice Hall.

Steinman, L. (2002). A touch of class: Considering the cloze. *The Canadian Modern Language Review, 59*, 291-301.

Storey, P. (1997). Examining the test-taking process: A cognitive perspective on the discourse cloze test. *Language Testing, 14*, 214-231.

Stubbs, J. B., & Tucker, G. R. (1974). The cloze test as a measure of English proficiency. *The Modern Language Journal, 58*, 239-241.

Taylor, W. L. (1953). Cloze procedure: A new tool for measuring readability. *Journalism Quarterly, 30*, 413-453.

Thongsa-nga, T. (1998). *A validation of the NC-Test for Mattayom six students.* Master's Thesis, The National Institute of Development Administration, Bangkok, Thailand.

Vacca, J. L., & Vacca, R. T. (2003). *Reading and learning to read* (5th ed.). Boston: Allyn and Bacon.

Weir, C. (1990). *Communicative language testing*. Hemel Hempstead: Prentice Hall International.

Wolter, B. (2002). Assessing proficiency through word associations: Is there still hope? *System, 30*, 315-329.

Wonghiransombat, P. (1998). *A validation of the MC-Test as a test of English language proficiency.* Master's thesis, The National Institute of Development Administration, Bangkok, Thailand.

Wongsothorn, A. (1996). *Testing and evaluation of English language learning*. Bangkok: Chulalongkorn University Press.

Yamashita, J. (2003). Processes of taking a gap-filling test: Comparison of skilled and less skilled EFL readers. *Language Testing, 20*, 267-293.

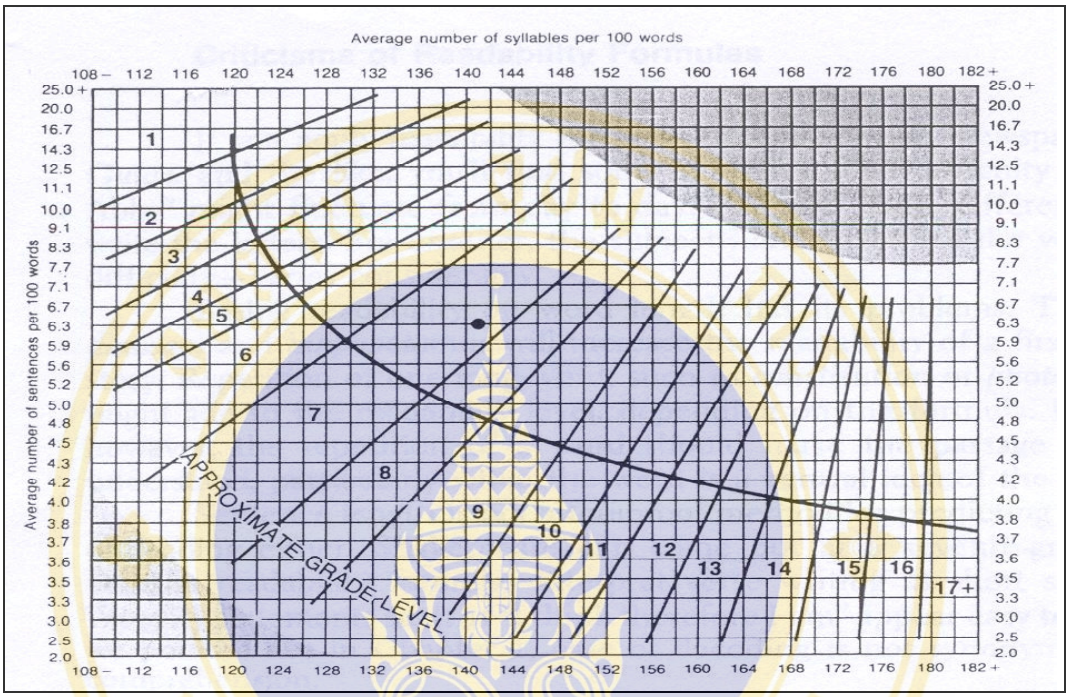
Copyright by Mahidol University

APPENDIX

Copyright by Mahidol University

APPENDIX A

FRY'S READABILITY GRAPH



Expanded Directions for Working Readability Graph

1. Randomly select three (3) sample passages and count out exactly 100 words each, beginning with the beginning of a sentence. Do count proper nouns, initializations, and numerals.
2. Count the number of sentences in the hundred words, estimating length of the fraction of the last sentence to the nearest one-tenth.
3. Count the total number of syllables in the 100-word passage. If you don't have a hand counter available, an easy way is to simply put a mark above every syllable over one in each word; then, when you get to the end of the passage, count the number of marks and add 100. Small calculators can also be used as counters by pushing numeral 1, then push the + sign for each word or syllable when counting.
4. Enter graph with average sentence length and average number of syllables; plot dot where the two lines intersect. Area where dot is plotted will give you the approximate grade level.
5. If a great deal of variability is found in syllable count or sentence count, putting more samples into the average is desirable
6. A word is defined as a group of symbols with a space on either side; thus, ***Joe, IRA 1945***, and ***&*** are each one word.
7. A syllable is defined as a phonetic syllable. Generally, there are as many syllables as vowel sounds. For example, ***stopped*** is one syllable and ***wanted*** is two syllables. When counting syllables for numerals and initializations, count one syllable for each symbol. For example, ***1945*** is four syllables, ***IRA*** is three syllables, and **&** is one syllable.

(Extracted from Fry, 1977, pp. 242-252)

Copyright by Mahidol University

APPENDIX B

THE C-TEST

DIRECTIONS:

In each of the six passages below, half of every second word in the passages has been taken out. Please fill in the missing part to complete each passage.

EXAMPLE:

THE C-TEST

PASSAGE A

Pesticides can help control pests. But **th** _ _ can **ca** _ _ _ harm, **t** _ _. People **m** _ _ not **fol** _ _ _ the **direc** _ _ _ _ _ carefully **wh** _ _ using **th** _ _.

ANSWERS:

Pesticides can help control pests. But **the y** can **cau s e** harm, **to o**. People **ma y** not **foll o w** the **direct i o n s** carefully **whe n** using **the m**.

Copyright by Mahidol University

แบบทดสอบนี้เป็นส่วนหนึ่งของงานวิจัยในระดับปริญญาโท
หลักสูตรภาษาศาสตร์ประยุกต์ ภาควิชาภาษาต่างประเทศ
คณะวิทยาศาสตร์ มหาวิทยาลัยมหิดล

คำชี้แจงในการทำแบบทดสอบ :

1. **แบบทดสอบนี้ประกอบด้วยบทความทั้งหมด 6 เรื่อง มี 4 หน้า ซึ่งครึ่งหลังของคำในทุก ๆ คำที่สองของแต่ละบทความจะถูกตัดออก**
2. **โปรดเติมส่วนที่ตัดออกไปของแต่ละคำในบทความทั้ง 6 เรื่องให้ครบถ้วน**
3. **โปรดเขียนคำตอบด้วยตัวบรรจง เพื่อความชัดเจนต่อการตรวจคำตอบ**

วัตถุประสงค์ :

ผู้วิจัยมีความประสงค์เป็นอย่างยิ่งที่จะขอความร่วมมือจากนักศึกษาให้ทำแบบทดสอบชุดนี้โดยสมบูรณ์ เพื่อนำไปใช้เป็นข้อมูลสำหรับงานวิทยานิพนธ์ในครั้งนี้

ตัวอย่าง :

THE C-TEST

PASSAGE A

Pesticides can help control pests. But **th** _ _ can **ca** _ _ _ harm, **t** _ _. People **m** _ _ not **fol** _ _ _ the **direc** _ _ _ _ _ carefully **wh** _ _ using **th** _ _.

คำตอบ :

Pesticides can help control pests. But **the y** can **cau s e** harm, **to o**. People **ma y** not **foll o w** the **direct i o n s** carefully **whe n** using **the m**.

Copyright by Mahidol University

คณะวิทยาศาสตร์ มหาวิทยาลัยมหิดล

วิชา ภาษาอังกฤษ **นักศึกษาปีที่ 1 SC**

วันที่ 2 มิถุนายน 2546 **เลขประจำตัว** _____

เวลา 15.30-16.30น. **กลุ่ม** _____

ชื่อ-นามสกุล (ไทย) _____ **อาจารย์ผู้สอน** _____

PASSGE ONE

A disease is a condition that impairs the proper functions of the body or of one of its part. Every **liv** _ _ _ thing, **bo** _ _ plants **a** _ _ animals, **c** _ _ provide **dis** _ _ _ _. People, **f** _ _ example, **a** _ _ often **infe** _ _ _ _ by **ti** _ _ bacteria. **A** _ _ bacteria, **i** _ turn, **c** _ _ be **infe** _ _ _ _ by **ev** _ _ smaller **vir** _ _ _ _. There **a** _ _ hundreds **o** _ different **dise** _ _ _ _. Each **h** _ _ its **o** _ _ particular **s** _ _ of **symp** _ _ _ _ or **si** _ _ _. These **a** _ _ clues **th** _ _ enable a doctor to diagnose the problem.

Copyright by Mahidol University

PASSAGE TWO

Human beings are a unique species. They **ha** _ _ more **com** _ _ _ _ brains **th** _ _ any **ot** _ _ _ beings. **Th** _ _ have **t** _ _ gift **o** _ speech **a** _ _ can **commu** _ _ _ _ _ _ their **thou** _ _ _ _, ideas, **a** _ _ feelings. **Hu** _ _ _ beings **a** _ _ interested **i** _ the **pa** _ _ and **c** _ _ learn **fr** _ _ past **exper** _ _ _ _ _ _ _. They **st** _ _ _ their **pre** _ _ _ _ _ problem **care** _ _ _ _ _ and **t** _ _ to **fi** _ _ the **be** _ _ solutions. **Th** _ _ look to the future and plan their next activities.

PASSAGE THREE

Do you smoke? If you do, you are one of more than a billion people in the world today who smoke. If **y** _ _ do **n** _ _ smoke, **y** _ _ are **sm** _ _ _, and **y** _ _ are **prob** _ _ _ _ healthier **th** _ _ many **smo** _ _ _ _. Cigarette **smo** _ _ _ _ is **com** _ _ _ all **ov** _ _ the **wo** _ _ _. It **i** _ also **ve** _ _ dangerous. **Smo** _ _ _ _ causes **ma** _ _ different **ki** _ _ _ of **ser** _ _ _ _ diseases, **su** _ _ as **lu** _ _ cancer **a** _ _ heart **dis** _ _ _ _. People **w** _ _ smoke **usu** _ _ _ _ know **th** _ _ the effects of smoking are very bad for them. So why do they smoke?

Copyright by Mahidol University

PASSAGE FOUR

Running for fitness, exercise, and pleasure is commonly called jogging. It **h** _ _ become **ve** _ _ popular **i** _ recent **ye** _ _ _. The **popul** _ _ _ _ _ of **jog** _ _ _ _ today **st** _ _ _ from **sev** _ _ _ _ factors. **Fi** _ _ _ jogging **i** _ one **o** _ the **mo** _ _ efficient **fo** _ _ _ of **exer** _ _ _ _. As _ rule, _ person **jog** _ _ _ _ burns **u** _ more **calo** _ _ _ _ per **min** _ _ _ than **i** _ most **ot** _ _ _ sports. **Run** _ _ _ _, like **bik** _ _ _, swimming, **a** _ _ quick walking, is an aerobic exercise

ขอขอบคุณนักศึกษาเป็นอย่างมากสำหรับการให้ความร่วมมือในการทำแบบทดสอบฉบับนี้ โดยสมบูรณ์

Copyright by Mahidol University

ANSWER KEY OF THE C-TEST

PASSAGE ONE

A disease is a condition that impairs the proper functions of the body or of one of its parts. Every 1)**liv i n g** thing, 2)**bo t h** plants 3)**a n d** animals, 4)**c a n** provide 5)**dis e a s e**. People, 6)**f o r** example, 7)**a r e** often 8)**infe c t e d** by 9)**ti n y** bacteria. 10)**A n d** bacteria, 11)**i n** turn, 12)**c a n** be 13)**infe c t e d** by 14)**ev e n** smaller 15)**vir u s e s**. There 16)**a r e** hundreds 17)**o f** different 18)**dise a s e s**. Each 19)**h a s** its 20)**o w n** particular 21)**s e t** of 22)**symp t o m s** or 23)**si g n s**. These 24)**a r e** clues 25)**th a t** enable a doctor to diagnose the problem.

Copyright by Mahidol University

PASSAGE TWO

Human beings are a unique species. They 26)**ha v e** more 27)**com p l e x** brains 28)**th a n** any 29)**ot h e r** beings. 30)**Th e y** have 31)**t h e** gift 32)**o f** speech 33)**a n d** can 34)**commu n i c a t e** their 35)**thou g h t s**, ideas, 36)**a n d** feelings. 37)**Hu m a n** beings 38)**a r e** interested 39)**i n** the 40)**pa s t** and 41)**c a n** learn 42)**fr o m** past 43)**exper i e n c e s**. They 44)**st u d y** their 45)**pre s e n t** problems 46)**care f u l l y** and 47)**t r y** to 48)**fi n d** the 49)**be s t** solutions. 50)**Th e y** look to the future and plan their next activities.

PASSAGE THREE

Do you smoke? If you do, you are one of more than a billion people in the world today who smoke. If 51)**y o u** do 52)**n o t** smoke, 53)**y o u** are 54)**sm a r t**, and 55)**y o u** are 56)**prob a b l y** healthier 57)**th a n** many 58)**smo k e r s**. Cigarette 59)**smo k i n g** is 60)**com m o n** all 61)**ov e r** the 62)**wo r l d**. It 63)**i s** also 64)**ve r y** dangerous. 65)**Smo k i n g** causes 66)**ma n y** different 67)**ki n d s** of 68)**ser i o u s** diseases, 69)**su c h** as 70)**lu n g** cancer 71)**a n d** heart 72)**dis e a s e**. People 73)**w h o** smoke 74)**usu a l l y** know 75)**th a t** the effects of smoking are very bad for them. So why do they smoke?

Copyright by Mahidol University

PASSAGE FOUR

Running for fitness, exercise, and pleasure is commonly called jogging. It 76)**h a s** become 77)**ve r y** popular 78)**i n** recent 79)**ye a r s**. The 80)**popul a r i t y** of 81)**jog g i n g** today 82)**st e m s** from 83)**sev e r a l** factors. 84)**Fi r s t** jogging 85)**i s** one 86)**o f** the 87)**mo s t** efficient 88)**fo r m s** of 89)**exer c i s e**. As 90)**a** rule, 91)**a** person 92)**jog g i n g** burns 93)**u p** more 94)**calo r i e s** per 95)**min u t e** than 96)**i n** most 97)**ot h e r** sports. 98)**Run n i n g**, like 99)**bik i n g**, swimming, 100)**a n d** quick walking, is an aerobic exercise.

Copyright by Mahidol University

APPENDIX C
THE NC-TEST

DIRECTIONS:

In each of the six passages below, half of every third word in the passages has been taken out. Please fill in the missing part to complete each passage.

EXAMPLE:

THE NEW C-TEST

PASSAGE A

Pesticides can help control pests. But they **c** _ _ cause harm, **t** _ _. People may **n** _ _ follow the **direc** _ _ _ _ _ carefully when **us** _ _ _ them.

ANSWERS:

Pesticides can help control pests. But they **ca n** cause harm, **to o**. People may **no t** follow the **direct i o n s** carefully when **usi n g** them.

Copyright by Mahidol University

แบบทดสอบนี้เป็นส่วนหนึ่งของงานวิจัยในระดับปริญญาโท
หลักสูตรภาษาศาสตร์ประยุกต์ ภาควิชาภาษาต่างประเทศ
คณะวิทยาศาสตร์ มหาวิทยาลัยมหิดล

คำชี้แจงในการทำแบบทดสอบ :

1. **แบบทดสอบนี้ประกอบด้วยบทความทั้งหมด 6 เรื่อง มี 4 หน้า ซึ่งครึ่งหลังของคำในทุก ๆ คำที่สาม ของแต่ละบทความจะถูกตัดออก**
2. **โปรดเติมส่วนที่ตัดออกไปของแต่ละคำในบทความทั้ง 6 เรื่องให้ครบถ้วน**
3. **โปรดเขียนคำตอบด้วยตัวบรรจง เพื่อความชัดเจนต่อการตรวจคำตอบ**

วัตถุประสงค์ :

ผู้วิจัยมีความประสงค์เป็นอย่างยิ่งที่จะขอความร่วมมือจากนักศึกษาให้ทำแบบทดสอบชุดนี้โดยสมบูรณ์ เพื่อนำไปใช้เป็นข้อมูลสำหรับงานวิทยานิพนธ์ในครั้งนี้

ตัวอย่าง :

THE NEW C-TEST

PASSAGE A

Pesticides can help control pests. But they **c _ _** cause harm, **t _ _**. People may **n _ _** follow the **direc _ _ _ _ _** carefully when **us _ _ _** them.

คำตอบ :

Pesticides can help control pests. But they **ca n** cause harm, **to o**. People may **no t** follow the **direct i o n s** carefully when **usi n g** them.

Copyright by Mahidol University

คณะวิทยาศาสตร์ มหาวิทยาลัยมหิดล

วิชา ภาษาอังกฤษ **นักศึกษาปีที่ 1 SC**

วันที่ 2 มิถุนายน 2546 **เลขประจำตัว** _____

เวลา 15.30-16.30 น. **กลุ่ม** _____

ชื่อ-นามสกุล (ไทย) _____ **อาจารย์ผู้สอน**

_____

PASSAGE ONE

A disease is a condition that impairs the proper functions of the body or of one of its part. Every living **th** _ _ _, both plants **a** _ _ _ animals, can **pro** _ _ _ _ disease. People, **f** _ _ example, are **of** _ _ _ infected by **ti** _ _ bacteria. And **bact** _ _ _ _, in turn, **c** _ _ be infected **b** _ even smaller **vir** _ _ _ _. There are **hund** _ _ _ _ of different **dise** _ _ _ _. Each has **i** _ _ own particular **s** _ _ of symptoms **o** _ signs. These **a** _ _ clues that **ena** _ _ _ a doctor **t** _ diagnose the **pro** _ _ _ _. A symptom **i** _ something a **pat** _ _ _ _ can detect, **su** _ _ as fever, **blee** _ _ _ _, or pain. _ sign is **some** _ _ _ _ _ a doctor can detect, such as swollen blood vessel or an enlarged internal body organ.

Copyright by Mahidol University

PASSAGE TWO

Human beings are a unique species. They have **mo** _ _ complex brains **th** _ _ any other **bei** _ _ _. They have **t** _ _ gift of **spe** _ _ _ and can **commu** _ _ _ _ _ _ their thoughts, **id** _ _ _, and feelings. **Hu** _ _ _ beings are **inter** _ _ _ _ _ in the **pa** _ _ and can **le** _ _ _ from past **exper** _ _ _ _ _ _. They study **th** _ _ _ present problems **care** _ _ _ _ _ and try **t** _ find the **be** _ _ solutions. They **lo** _ _ to the **fut** _ _ _ and plan **th** _ _ _ next activities. **Th** _ _ invent tools **a** _ _ machines to **he** _ _ them in **th** _ _ _ work. Can **mach** _ _ _ _ perform in **th** _ _ _ ways? Some computer scientists are developing machines that may think and act in some ways like human beings.

Copyright by Mahidol University

PASSAGE THREE

Do you smoke? If you do, you are one of more than a billion people in the world today who smoke. If you **d** _ not smoke **y** _ _ are smart, **a** _ _ you are **prob** _ _ _ _ healthier than **ma** _ _ smokers. Cigarette **smo** _ _ _ _ is common **a** _ _ over the **wo** _ _ _. It is **al** _ _ very dangerous. **Smo** _ _ _ _ causes many **diff** _ _ _ _ _ kinds of **ser** _ _ _ _ diseases, such **a** _ lung cancer **a** _ _ heart disease. **Peo** _ _ _ who smoke **usu** _ _ _ _ know that **t** _ _ effects of **smo** _ _ _ _ are very **b** _ _ for them. **S** _ why do **th** _ _ smoke? Smoking **i** _ a habit, _ bad habit. **Ma** _ _ people start **t** _ smoke when they are teenagers, usually because they have friends who smoke and they want to be like them.

Copyright by Mahidol University

PASSAGE FOUR

Running for fitness, exercise, and pleasure is commonly called jogging. It has **bec** _ _ _ very popular **i** _ recent years. **T** _ _ popularity of **jog** _ _ _ _ today stems **fr** _ _ several factors. **Fi** _ _ _ jogging is **o** _ _ of the **mo** _ _ efficient forms **o** _ exercise. As _ rule, a **per** _ _ _ jogging burns **u** _ more calories **p** _ _ minute than **i** _ most other **spo** _ _ _. Running, like **bik** _ _ _, swimming, and **qu** _ _ _ walking, is **a** _ aerobic exercise. **Su** _ _ an exercise **us** _ _ a great **de** _ _ of oxygen. **I** _ addition, it **incr** _ _ _ _ _ the heart **ra** _ _. Aerobic exercise **stren** _ _ _ _ _ _ the heart muscle so that it pumps more efficiently.

ขอขอบคุณนักศึกษาเป็นอย่างมากสำหรับการให้ความร่วมมือในการทำแบบทดสอบฉบับนี้ โดยสมบูรณ์

Copyright by Mahidol University

ANSWER KEY OF THE NC-TEST

PASSAGE ONE

A disease is a condition that impairs the proper functions of the body or of one of its parts. Every living 1)**th i n g**, both plants 2)**a n d** animals, can 3)**pro v i d e** disease. People, 4)**f o r** example, are 5)**of t e n** infected by 6)**ti n y** bacteria. And 7)**bact e r i a**, in turn, 8)**c a n** be infected 9)**b y** even smaller 10)**vir u s e s**. There are 11)**hund r e d s** of different 12)**dise a s e s**. Each has 13)**i t s** own particular 14)**s e t** of symptoms 15)**o r** signs. These 16)**a r e** clues that 17)**ena b l e** a doctor 18)**t o** diagnose the 19)**pro b l e m**. A symptom 20)**i s** something a 21)**pat i e n t** can detect, 22)**su c h** as fever, 23)**blee d i n g**, or pain. 24)**A** sign is 25)**some t h i n g** a doctor can detect, such as swollen blood vessel or an enlarged internal body organ.

Copyright by Mahidol University

PASSAGE TWO

Human beings are a unique species. They have 26)**mo r e** complex brains 27)**th a n** any other 28)**bei n g s**. They have 29)**t h e** gift of 30)**spe e c h** and can 31)**commu n i c a t e** their thoughts, 32)**id e a s**, and feelings. 33)**Hu m a n** beings are 34)**inter e s t e d** in the 35)**pa s t** and can 36)**le a r n** from past 37)**exper i e n c e s**. They study 38)**th e i r** present problems 39)**care f u l l y** and try 40)**t o** find the 41)**be s t** solutions. They 42)**lo o k** to the 43)**fut u r e** and plan 44)**th e i r** next activities. 45)**Th e y** invent tools 46)**a n d** machines to 47)**he l p** them in 48)**th e i r** work. Can 49)**mach i n e s** perform in 50)**th e s e** ways? Some computer scientists are developing machines that may think and act in some ways like human beings.

Copyright by Mahidol University

PASSAGE THREE

Do you smoke? If you do, you are one of more than a billion people in the world today who smoke. If you 51)**d o** not smoke, 52)**y o u** are smart, 53)**a n d** you are 54)**prob a b l y** healthier than 55)**ma n y** smokers. Cigarette 56)**smo k i n g** is common 57)**a l l** over the 58)**wo r l d**. It is 59)**al s o** very dangerous. 60)**Smo k i n g** causes many 61)**diff e r e n t** kinds of 62)**ser i o u s** diseases, such 63)**a s** lung cancer 64)**a n d** heart disease. 65)**Peo p l e** who smoke 66)**usu a l l y** know that 67)**t h e** effects of 68)**smo k i n g** are very 69)**b a d** for them. 70)**S o** why do 71)**th e y** smoke? Smoking 72)**i s** a habit, 73)**a** bad habit. 74)**Ma n y** people start 75)**t o** smoke when they are teenagers, usually because they have friends who smoke and they want to be like them.

Copyright by Mahidol University

PASSAGE FOUR

Running for fitness, exercise, and pleasure is commonly called jogging. It has 76)**bec o m e** very popular 77)**i n** recent years. 78)**T h e** popularity of 79)**jog g i n g** today stems 80)**fr o m** several factors. 81)**Fi r s t** jogging is 82)**o n e** of the 83)**mo s t** efficient forms 84)**o f** exercise. As 85)**a** rule, a 86)**per s o n** jogging burns 87)**u p** more calories 88)**p e r** minute than 89)**i n** most other90)**spo r t s**. Running, like 91)**bik i n g**, swimming, and 92)**qu i c k** walking, is 93)**a n** aerobic exercise. 94)**Su c h** an exercise 95)**us e s** a great96)**de a l** of oxygen.97)**I n** addition, it 98)**incr e a s e s** the heart 99)**ra t e**. Aerobic exercise 100)**stren g t h e n s** the heart muscle so that it pumps more efficiently.

Copyright by Mahidol University

APPENDIX D
THE MC-TEST

DIRECTIONS:

In each of the six passages below, half of every second word in the passages has been taken out. Please fill in the missing part to complete each passage.

EXAMPLE:

THE MODIFIED C-TEST

PASSAGE A

Pesticides can help control pests. But _ _ **ey** can _ _ _ **se** harm,_ _ **o**. People _ _ **y** not _ _ _ **low** the _ _ _ _ _ **tions** carefully _ _ **en** using _ _ **em**.

ANSWERS:

Pesticides can help control pests. But **t hey** can **c a use** harm, **t oo**. People **m ay** not **f o llow** the **d i r e ctions** carefully **w hen** using **t h**em

Copyright by Mahidol University

แบบทดสอบนี้เป็นส่วนหนึ่งของงานวิจัยในระดับปริญญาโท
หลักสูตรภาษาศาสตร์ประยุกต์ ภาควิชาภาษาต่างประเทศ
คณะวิทยาศาสตร์ มหาวิทยาลัยมหิดล

คำชี้แจงในการทำแบบทดสอบ :

1. **แบบทดสอบนี้ประกอบด้วยบทความทั้งหมด 6 เรื่อง มี 4 หน้า ซึ่งครึ่งหลังของคำในทุก ๆ คำที่สองของแต่ละบทความจะถูกตัดออก**
2. **โปรดเติมส่วนที่ตัดออกไปของแต่ละคำในบทความทั้ง 6 เรื่องให้ครบถ้วน**
3. **โปรดเขียนคำตอบด้วยตัวบรรจง เพื่อความชัดเจนต่อการตรวจคำตอบ**

วัตถุประสงค์ :

ผู้วิจัยมีความประสงค์เป็นอย่างยิ่งที่จะขอความร่วมมือจากนักศึกษาให้ทำแบบทดสอบชุดนี้โดยสมบูรณ์ เพื่อนำไปใช้เป็นข้อมูลสำหรับงานวิทยานิพนธ์ในครั้งนี้

ตัวอย่าง :

THE MODIFIED C-TEST

PASSAGE A

Pesticides can help control pests. But _ _ **ey** can _ _ _ **se** harm,_ _ **o**. People _ _ **y** not _ _ _ **low** the _ _ _ _ _ **tions** carefully _ _ **en** using _ _ **em**.

คำตอบ :

Pesticides can help control pests. But **the y** can **cau s e** harm, **to o**. People **ma y** not **foll o w** the **direct i o n s** carefully **whe n** using **the m**.

Copyright by Mahidol University

คณะวิทยาศาสตร์ มหาวิทยาลัยมหิดล

วิชา ภาษาอังกฤษ **นักศึกษาปีที่ 1 SC**

วันที่ 2 มิถุนายน 2546 **เลขประจำตัว** _____

เวลา 15.30-16.30น. **กลุ่ม** _____

ชื่อ-นามสกุล (ไทย) _____ **อาจารย์ผู้สอน** _____

PASSGE ONE

A disease is a condition that impairs the proper functions of the body or of one of its part. Every _ _ _ **ing** thing, _ _ **th** plants _ _ **d** animals, _ _ **n** provide _ _ _ _ **ase**. People, _ _ **r** example, _ _ **e** often _ _ _ _ **cted** by _ _ **ny** bacteria. _ _ **d** bacteria, _ **n** turn, _ _ **n** be _ _ _ _ **cted** by _ _ **en** smaller _ _ _ _ **ses**. There _ _ **e** hundreds _ **f** different _ _ _ _ **ases**. Each _ _ **s** its _ _ **n** particular _ _ **t** of _ _ _ _ **toms** or _ _ _ **ns**. These _ _ **e** clues _ _ **at** enable a doctor to diagnose the problem.

Copyright by Mahidol University

PASSAGE TWO

Human beings are a unique species. They _ _ **ve** more _ _ _ _ **lex** brains _ _ **an** any _ _ _ **er** beings. _ _ **ey** have _ _ **e** gift _ **f** speech _ _ **d** can _ _ _ _ _ _ **icate** their _ _ _ _ **ghts**, ideas, _ _ **d** feelings. _ _ _ **an** beings _ _ **e** interested _ **n** the _ _ **st** and _ _ **n** learn _ _ **om** past _ _ _ _ _ _ **ences**. They _ _ _ **dy** their _ _ _ **sent** problems _ _ _ _ _ **ully** and _ _ **y** to _ _ **nd** the _ _ **st** solutions. look to the future and plan their next activities

PASSAGE THREE

Do you smoke? If you do, you are one of more than a billion people in the world today who smoke. If _ _ **u** do _ _ **t** smoke, _ _ **u** are _ _ _ **rt**, and _ _ **u** are _ _ _ _ **ably** healthier _ _ **an** many _ _ _ _ **ers** . Cigarette _ _ _ _ **ing** is _ _ _ **mon** all _ _ **er** the _ _ _ **ld**. It _ **s** also _ _ **ry** dangerous. _ _ _ _ **ing** causes _ _ **ny** different _ _ _ **ds** of _ _ _ _ **ous** diseases, _ _ **ch** as _ _ **ng** cancer _ _ **d** heart _ _ _ _ **ase**. People _ _ **o** smoke _ _ _ _ **lly** know _ _ **at** the effects of smoking are very bad for them. So why do they smoke?

Copyright by Mahidol University

PASSAGE FOUR

Running for fitness, exercise, and pleasure is commonly called jogging. It _ _ **s** become _ _ **ry** popular _ **n** recent _ _ _ **rs**. The _ _ _ _ _ **arity** of _ _ _ _ **ing** today _ _ _ **ms** from _ _ _ _ **ral** factors. _ _ _ **st** jogging _ **s** one _ **f** the _ _ **st** efficient _ _ _ **ms** of _ _ _ _ **cise**. As _ rule, _ person _ _ _ _ **ing** burns _ **p** more _ _ _ _ **ries** per _ _ _ **ute** than _ **n** most _ _ _ **er** sports. _ _ _ _ **ing**, like _ _ _ **ing**, swimming, _ _ **d** quick walking, is an aerobic exercise.

ขอขอบคุณนักศึกษาเป็นอย่างมากสำหรับการให้ความร่วมมือในการทำแบบทดสอบฉบับนี้ โดยสมบูรณ์

Copyright by Mahidol University

ANSWER KEY OF THE MC-TEST

PASSAGE ONE

A disease is a condition that impairs the proper functions of the body or of one of its parts. Every 1)**l i v ing** thing, 2)**b o th** plants 3)**a n d** animals, 4)**c a n** provide 5)**d i s e ase**. People, 6)**f o r** example, 7)**a r e** often 8)**i n f e c ted** by 9)**t i ny** bacteria. 10)**A n d** bacteria, 11)**i n** turn, 12)**c a n** be 13)**i n f e cted** by 14)**e v en** smaller 15)**v i r u ses**. There 16)**a r e** hundreds 17)**o f** different 18)**d i s e ases**. Each 19)**h a s** its 20)**o w n** particular 21)**s e t** of 22)**s y m p toms** or 23)**s i g ns**. These 24)**a r e** clues 25)**t h at** enable a doctor to diagnose the problem.

Copyright by Mahidol University

PASSAGE TWO

Human beings are a unique species. They 26)**h a ve** more 27)**c o m p lex** brains 28)**t h an** any 29)**o t h er** beings. 30)**T h ey** have 31)**t h e** gift 32)**o f** speech 33)**a n d** can 34) **c o m m u n icate** their 35)**t h o u ghts**, ideas, 36)**a n d** feelings. 37)**H u m an** beings 38)**a r e** interested 39)**i n** the 40)**p a st** and 41)**c a n** learn 42)**f r om** past 43)**e x p e r i ences**. They 44)**s t u dy** their 45)**p r e sent** problems 46)**c a r e f ully** and 47)**t r y** to 48)**f i nd** the 49)**b e st** solutions.50)**T h ey** look to the future and plan their next activities.

PASSAGE THREE

Do you smoke? If you do, you are one of more than a billion people in the world today who smoke. If 51)**y o u** do 52)**n o t** smoke, 53)**y o u** are 54)**s m a rt**, and 55)**y o u** are 56)**p r o b ably** healthier 57)**t h an** many 58)**s m o k ers**. Cigarette 59)**s m o k ing** is 60)**c o m mon** all 61)**o v er** the 62)**w o r ld**. It 63)**i s** also 64)**v e ry** dangerous. 65)**S m o k ing** causes 66)**m a ny** different 67)**k i n ds** of 68)**s e r i ous** diseases,69)**s u ch** as 70)**l u ng** cancer 71)**a n d** heart 72)**d i s e ase**. People 73)**w h o** smoke 74)**u s u a lly** know 75)**t h at** the effects of smoking are very bad for them. So why do they smoke?

Copyright by Mahidol University

PASSAGE FOUR

Running for fitness, exercise, and pleasure is commonly called jogging. It 76)**h a s** become 77)**v e ry** popular 78)**i n** recent 79)**y e a rs**. The 80)**p o p u l arity** of 81)**j o g g ing** today (82)**ste m s** from 83)**s e v e ral** factors. 84)**F i r st** jogging 85)**i s** one 86)**o f** the 87)**m o st** efficient 88)**f o r ms** of 89)**e x e r cise**. As 90)**a** rule, 91)**a** person 92)**j o g g ing** burns 93)**u p** more 94)**c a l o ries** per 95)**m i n ute** than 96)**i n** most 97)**o t h er** sports. 98)**R u n n ing**, like 99)**b i k ing**, swimming, 100)**a n d** quick walking, is an aerobic exercise.

Copyright by Mahidol University

APPENDIX E
THE NEW MODIFIED CLOZE-TESTS

DIRECTIONS:

In each of the six passages below, half of every third word in the passages has been taken out. Please fill in the missing part to complete each passage.

EXAMPLE:

THE NEW MODIFIED C-TEST

PASSAGE A

Pesticides can help control pests. But they _ _ **n** cause harm, _ _ **o**. People may _ _ **t** follow the _ _ _ _ _ **tions** carefully when _ _ _ **ng** them.

ANSWERS:

Pesticides can help control pests. But they **c an** cause harm, **t o o**. People may **n ot** follow the **d i r e ctions** carefully when **u s ing** them.

Copyright by Mahidol University

แบบทดสอบนี้เป็นส่วนหนึ่งของงานวิจัยในระดับปริญญาโท
หลักสูตรภาษาศาสตร์ประยุกต์ ภาควิชาภาษาต่างประเทศ
คณะวิทยาศาสตร์ มหาวิทยาลัยมหิดล

คำชี้แจงในการทำแบบทดสอบ :

1. **แบบทดสอบนี้ประกอบด้วยบทความ 6 เรื่อง ทั้งหมดมี 4 หน้า ซึ่งครึ่งแรกของคำในทุกๆ คำที่สามของบทความจะถูกตัดออก**
2. **โปรดเติมส่วนที่ตัดออกไปของแต่ละคำในบทความทั้ง 6 เรื่องให้ครบถ้วน**
3. **โปรดเขียนคำตอบด้วยตัวบรรจง เพื่อความชัดเจนต่อการตรวจคำตอบ**

วัตถุประสงค์ :

ผู้วิจัยมีความประสงค์เป็นอย่างยิ่งที่จะขอความร่วมมือจากนักศึกษาให้ทำแบบทดสอบชุดนี้โดยสมบูรณ์ เพื่อนำไปใช้เป็นข้อมูลสำหรับงานวิทยานิพนธ์ในครั้งนี้

ตัวอย่าง :

THE NEW MODIFIED CLOZE TEST

PASSAGE A

Pesticides can help control pests. But they _ _ **n** cause harm, _ _ **o**. People may _ _ **t** follow the _ _ _ _ _ **tions** carefully when _ _ _ **ng** them.

คำตอบ :

Pesticides can help control pests. But they **c an** cause harm, **t o o**. People may **n ot** follow the **d i r e ctions** carefully when **u s ing** them.

Copyright by Mahidol University

คณะวิทยาศาสตร์ มหาวิทยาลัยมหิดล

วิชา ภาษาอังกฤษ **นักศึกษาปีที่ 1 SC**

วันที่ 2 มิถุนายน 2546 **เลขประจำตัว** _____

เวลา 15.30-16.30 น. **กลุ่ม** _____

ชื่อ-นามสกุล (ไทย) _____ **อาจารย์ผู้สอน** _____

PASSAGE ONE

A disease is a condition that impairs the proper functions of the body or of one of its part. Every living _ _ _ **ng**, both plants _ _ **d** animals, can _ _ _ _ **ide** disease. People, _ _ **r** example, are _ _ _ **en** infected by _ _ **ny** bacteria. And _ _ _ _ **eria**, in turn, _ _ **n** be infected _ **y** even smaller _ _ _ _ **ses**. There are _ _ _ _ **reds** of different _ _ _ _ **ases**. Each has _ _ **s** own particular _ _ **t** of symptoms _ **r** signs. These _ _ **e** clues that _ _ _ **ble** a doctor _ **o** diagnose the _ _ _ _ **lem**. A symptom _ **s** something a _ _ _ _ **ent** can detect, _ _ **ch** as fever, _ _ _ _ **ding**, or pain. _ sign is _ _ _ _ _ **hing** a doctor can detect, such as swollen blood vessel or an enlarged internal body organ.

Copyright by Mahidol University

PASSAGE TWO

Human beings are a unique species. They have _ _ **re** complex brains _ _ **an** any other _ _ _ **ngs**. They have _ _ **e** gift of _ _ _ **ech** and can _ _ _ _ _ _ **icate** their thoughts, _ _ _ **as**, and feelings. _ _ _ **an** beings are _ _ _ _ _ **ested** in the _ _ **st** and can _ _ _ **rn** from past _ _ _ _ _ _ **ences**. They study _ _ _ **ir** present problems _ _ _ _ _ **ully** and try _ **o** find the _ _ **st** solutions. They _ _ **ok** to the _ _ _ **ure** and plan _ _ _ **ir** next activities. _ _ **ey** invent tools _ _ **d** machines to _ _ **lp** them in _ _ _ **ir** work. Can _ _ _ _ **ines** perform in _ _ _ **se** ways? Some computer scientists are developing machines that may think and act in some ways like human beings.

Copyright by Mahidol University

PASSAGE THREE

Do you smoke? If you do, you are one of more than a billion people in the world today who smoke. If you _ **o** not smoke _ _ **u** are smart, _ _ **d** you are _ _ _ _ **ably** healthier than _ _ **ny** smokers. Cigarette _ _ _ _ **ing** is common _ _ **l** over the _ _ _ **ld**. It is _ _ _ **ays** very dangerous.

_ _ _ _ **ing** causes many _ _ _ _ _ _ **rent** kinds of _ _ _ _ **ous** diseases, such _ **s** lung cancer _ _ **d** heart disease.

_ _ _ **ple** who smoke _ _ _ _ **lly** know that _ _ **e** effects of _ _ _ _ **ing** are very _ _ **d** for them. _ **o** why do _ _ **ey** smoke? Smoking _ **s** a habit, _ bad habit. _ _ **ny** people start _ **o** smoke when they are teenagers, usually because they have friends who smoke and they want to be like them.

Copyright by Mahidol University

PASSAGE FOUR

Running for fitness, exercise, and pleasure is commonly called jogging. It has _ _ _ **ome** very popular _ **n** recent years. _ _ **e** popularity of _ _ _ _ **ing** today stems _ _ **om** several factors. _ _ _ **st** jogging is _ _ **e** of the _ _ **st** efficient forms _ **f** exercise. As _ rule, a _ _ _ **son** jogging burns _ **p** more calories _ _ **r** minute than _ **n** most other _ _ _ **rts**. Running, like _ _ _ **ing**, swimming, and _ _ _ **ck** walking, is _ **n** aerobic exercise. _ _ **ch** an exercise _ _ **es** a great _ _ **al** of oxygen. _ **n** addition, it _ _ _ _ _ **ases** the heart _ _ **te**. Aerobic exercise _ _ _ _ _ _ **thens** the heart muscle so that it pumps more efficiently

ขอขอบคุณนักศึกษาเป็นอย่างมากสำหรับการให้ความร่วมมือในการทำแบบทดสอบฉบับนี้ โดยสมบูรณ์

Copyright by Mahidol University

ANSWER KEY OF THE NEW MODIFIED C-TEST

PASSAGE ONE

A disease is a condition that impairs the proper functions of the body or of one of its parts. Every living 1)**t h i ng**, both plants 2)**a n d** animals, can 3)**p r o v ide** disease. People, 4)**f o r** example, are 5)**o f t en** infected by 6)**t i ny** bacteria. And 7)**b a c t eria**, in turn, 8)**c a n** be infected 9)**b y** even smaller 10)**v i r u ses**. There are 11)**h u n d reds** of different 12)**d i s e ases**. Each has 13)**i t s** own particular 14)**s e t** of symptoms 15)**o r** signs. These 16)**a r e** clues that 17)**e n a ble** a doctor 18)**t o** diagnose the 19)**p r o b lem**. A symptom 20)**i s** something a 21)**p a t i ent** can detect, 22)**s u ch** as fever, 23)**b l e e ding**, or pain. 24)**A** sign is 25)**s o m e t hing** a doctor can detect, such as swollen blood vessel or an enlarged internal body organ.

Copyright by Mahidol University

PASSAGE TWO

Human beings are a unique species. They have 26)**m o re** complex brains 27)**t h a n** any other 28)**b e i ngs.** They have 29)**t h e** gift of 30)**s p e ech** and can 31)**c o m m u n icate** their thoughts,32)**i d e as,** and feelings. 33)**H u m an** beings 34)**i n t e r ested** in the 35)**p a st** and can 36)**l e a rn** from past 37)**e x p e r i ences.** They study 38)**t h e ir** present problems 39)**c a r e f ully** and try 40)**t o** find the 41)**b e st** solutions. They 42)**l o ok** to the 43)**f u t ure** and plan 44)**t h e ir** next activities.45) **T h ey** invent tools 46)**a n d** machines to47)**h e lp** them in 48)**t h e ir** work. Can 49)**mach i n e s** perform in 50)**th e s e** ways? Some computer scientists are developing machines that may think and act in some ways like human beings.

Copyright by Mahidol University

PASSAGE THREE

Do you smoke? If you do, you are one of more than a billion people in the world today who smoke. If you 51)**d o** not smoke,52)**y o u** are smart, 53)**a n d** you are 54)**p r o b ably** healthier than 55)**m a ny** smokers. Cigarette 56)**s m o k ing** is common 57)**a l l** over the 58)**w o r l d**. It is 59)**a l so** very dangerous. 60)**S m o k i n g** causes many 61)**d i f f e rent** kinds of 62)**s e r i ous** diseases, such 63)**a s** lung cancer 64)**a n d** heart disease. 65)**P e o ple** who smoke 66)**u s u a lly** know that67)**t h e** effects of 68)**s m o k ing** are very 69)**b a d** for them. 70)**S o** why do 71)**t h ey** smoke? Smoking 72)**i s** a habit, 73)**a** bad habit. 74)**M a ny** people start 75)**t o** smoke when they are teenagers, usually because they have friends who smoke and they want to be like them.

Copyright by Mahidol University

PASSAGE FOUR

Running for fitness, exercise, and pleasure is commonly called jogging. It has 76)**b e c ome** very popular 77)**i n** recent years. 78)**T h e** popularity of 79)**j o g g ing** today stems 80)**f r om** several factors. 81)**F i r st** jogging is 82)**o n e** of the 83)**m o st** efficient forms 84)**o f** exercise. As 85)**a** rule, a 86)**p e r son** jogging burns 87)**u p** more calories 88)**p e r** minute than 89)**i n** most other 90)**s p o rts**. Running, like 91)**b i k ing**, swimming, and 92)**q u i c k** walking, is 93)**a n** aerobic exercise. 94)**S u ch** an exercise 95)**u s es** a great 96)**d e al** of oxygen. 97)**I n** addition, it 98)**i n c r e ases** the heart 99)**r a te**. Aerobic exercise100)**s t r e n g thens** the heart muscle so that it pumps more efficiently.

Copyright by Mahidol University

APPENDIX F

KUDER-RICHARDSON FORMULA 21 (K-R 21)

$$\text{K-R21} = \frac{k}{k-1}\left(1 - \frac{\bar{X}(k-\bar{X})}{kS^2}\right)$$

where K-R21= Kuder-Richardson Formula 21

k = number of items

$\bar{X}$ = means of the test scores

S = standard deviation of the test scores

1. The mean scores and the reliability of the four cloze tests

Type	Means	Std.	Reliability
The MC-Test	69.50	17.60	.9285
The NMC-Test	73.50	19.09	.9417
The C-Test	80.75	23.16	.9648
The NC-Test	69.50	19.62	.9470

2. The reliability of the four cloze tests by using K-R 21

MC-TEST K-R 21 $= \frac{k}{k-1}\left(1 - \frac{\bar{X}(k-\bar{X})}{kS^2}\right)$

$= \frac{100}{100-1}\left(1 - \frac{48.32\,(100-48.32)}{100 \times 17.60^2}\right)$

$= 1.01\,(1 - \frac{2497.17}{30976})$ $= 1.01\,(1 - .0806)$

$= 1.01 \times .9194$ $= \mathbf{.9285}$

Copyright by Mahidol University

NMC-TEST K-R 21 $= \frac{k}{k-1}\left(1 - \frac{\bar{X}(k-\bar{X})}{kS^2}\right)$

$= \frac{100}{100-1}\left(1 - \frac{55.78\,(100-55.78)}{100 \times 19.09^2}\right)$

$= 1.01\,(1 - \frac{2466.59}{36442}) = 1.01\,(1 - .0676)$

$= 1.01 \times .9324 = \mathbf{.9417}$

C-TEST K-R 21 $= \frac{k}{k-1}\left(1 - \frac{\bar{X}(k-\bar{X})}{kS^2}\right)$

$= \frac{100}{100-1}\left(1 - \frac{62.64\,(100-60.24)}{100 \times 23.16^2}\right)$

$= 1.01\,(1 - \frac{2395.14}{53638.56}) = 1.01\,(1 - .0446)$

$= 1.01 \times .9553 = \mathbf{.9648}$

NC-TEST K-R 21 $= \frac{k}{k-1}\left(1 - \frac{\bar{X}(k-\bar{X})}{kS^2}\right)$

$= \frac{100}{100-1}\left(1 - \frac{59.96\,(100-59.96)}{100 \times 19.62^2}\right)$

$= 1.01\,(1 - \frac{2400.79}{38494.44}) = 1.01\,(1 - .0623)$

$= 1.01 \times .9377 = \mathbf{.9470}$

Copyright by Mahidol University

3. The scores of the four cloze tests in the pilot study

Student	Type	Score	Means	Std.	Student	Type	Score	Means	Std.
1	MC-Test	75	69.50	10.96	**9**	NC-Test	86	69.50	12.39
2	MC-Test	82			**10**	NC-Test	72		
3	MC-Test	63			**11**	NC-Test	61		
4	MC-Test	58			**12**	NC-Test	59		
5	NMC-Test	80	73.50	10.40	**13**	C-Test	97	80.75	13.37
6	NMC-Test	59			**14**	C-Test	65		
7	NMC-Test	73			**15**	C-Test	77		
8	NMC-Test	82			**16**	C-Test	84		

4. The reliability of the four cloze tests in the pilot study by using K-R 21

$$\textbf{MC-TEST K-R 21} = \frac{k}{k-1}\left(1-\frac{\bar{X}(k-\bar{X})}{kS^2}\right)$$

$$= \frac{100}{100-1}\left(1-\frac{69.5\,(100-69.5)}{100\times 10.96^2}\right)$$

$$= 1.01\left(1-\frac{2119.75}{12012.16}\right) \qquad = 1.01\,(1-.1764)$$

$$= 1.01\times .8236 \qquad = \textbf{.8318}$$

$$\textbf{NMC-TEST K-R 21} = \frac{k}{k-1}\left(1-\frac{\bar{X}(k-\bar{X})}{kS^2}\right)$$

$$= \frac{100}{100-1}\left(1-\frac{73.50\,(100-73.50)}{100\times 10.40^2}\right)$$

$$= 1.01\left(1-\frac{1947.75}{10816}\right) \qquad = 1.01\,(1-.1800)$$

$$= 1.01\times .82 \qquad = \textbf{.8282}$$

Copyright by Mahidol University

C-TEST K-R 21 $= \frac{k}{k-1}\left(1 - \frac{\bar{X}(k-\bar{X})}{kS^2}\right)$

$= \frac{100}{100-1}\left(1 - \frac{80.75\,(100-80.75)}{100 \times 13.37^2}\right)$

$= 1.01\,(1 - \frac{1554.43}{17875.69})$ $= 1.01\,(1 - .0869)$

$= 1.01 \times .9131$ $= \mathbf{.9222}$

NC-TEST K-R 21 $= \frac{k}{k-1}\left(1 - \frac{\bar{X}(k-\bar{X})}{kS^2}\right)$

$= \frac{100}{100-1}\left(1 - \frac{69.50(100-69.50)}{100 \times 12.39^2}\right)$

$= 1.01\,(1 - \frac{2119.75}{15351.21})$ $= 1.01\,(1 - .1380)$

$= 1.01 \times .862$ $= \mathbf{.8706}$

Copyright by Mahidol University

APPENDIX G
INTERVIEW FORMAT

แบบฟอร์มการสัมภาษณ์นักศึกษา SC ปีที่ 1

1. ชื่อ: _____ **รหัสนักศึกษา:** _____

2. กลุ่ม SC: ☐ กลุ่มที่ 1 ☐ กลุ่มที่ 2 ☐ กลุ่มที่ 3 ☐ กลุ่มที่ 4 ☐ กลุ่มที่ 5

3. ประเภทแบบทดสอบ: ☐ MC-Test ☐ NMC-Test ☐ NC-Test ☐ C-Test

4. ระยะเวลาการสัมภาษณ์จะใช้เวลา 30-40 นาที โดยการเก็บข้อมูลการสัมภาษณ์จะใช้วิธี

4.1 จดบันทึกระหว่างการสัมภาษณ์

4.2 บันทึกเทประหว่างการสัมภาษณ์

5. คำถามที่ใช้ประกอบการสัมภาษณ์

5.1 กรุณาจัดลำดับความยากง่ายของแบบทดสอบ cloze ที่ทำไปแล้ว โดยใส่หมายเลขลงในช่องว่าง ดังนี้โดยหมายเลข (1) ง่ายที่สุด ↔ (4) ยากที่สุด

Passage 1	Passage 2	Passage 3	Passage 4

5.2 ขณะที่นักศึกษากำลังคิดหาคำตอบเติมลงไปในช่องว่างของแบบทดสอบที่ทำไปแล้ว นักศึกษาใช้วิธีการอะไรบ้างในการหาคำตอบมาเติมลงไปในแต่ละช่องว่าง

5.3 เมื่อเริ่มทำแบบทดสอบ นักศึกษาปฏิบัติดังนี้ (เลือกตอบได้มากกว่า 1 ข้อ)

ก. อ่านเนื้อเรื่องตลอดตั้งแต่ต้นจนจบก่อนแล้วจึงค่อยคิดหาคำตอบ

ข. อ่านเฉพาะบางส่วนของเนื้อความ

ค. อ่านกลับไปกลับมา เพื่อช่วยในการหาคำตอบ

5.4 ถ้านักศึกษาหาคำตอบไม่ได้ นักศึกษาใช้วิธีการเดาอย่างเดียว หรือหาทางใช้วิธีอื่นมาช่วย โปรดอธิบาย

Copyright by Mahidol University

INTERVIEW FORMAT OF CLOZE TEST–TAKING STRATEGIES

1. Name: _____ Student ID: _____

2. SC Group: □ group 1 □ group 2 □ group 3 □ group 4 □ group 5

3. Test Type: □ MC-Test □ NMC-Test □ NC-Test □ C-Test

4. The interview duration will take 30-40 minutes and the data will be collected by

- Note taking

- Tape Recording

5. Interview Questions:

5.1 Could you please write down the number of difficulty level for each passage? (Number 1 the easiest → 4 the most difficult)"

Passage 1	Passage 2	Passage 3	Passage 4

5.2 What procedures do you use while answering the new forms of the cloze tests?

5.3 Do you read from the beginning to the end, or only parts of the test, or jump around?

5.4 Do you use other ways to find the answer in the cloze tests?

Copyright by Mahidol University

APPENDIX H

ITEM ANALYSIS OF THE FOUR CLOZE TESTS

Item analysis of the MC-Test (part I)

Item Analysis of the MC-Test									
Item	**IF Total**	**IF Upper**	**IF Lower**	**ID**	**Item**	**IF Total**	**IF Upper**	**IF Lower**	**ID**
1	0.28	0.50	0.07	0.43	**26**	*0.85*	1.00	0.71	*0.29*
2	0.32	0.35	0.28	*0.07*	**27**	0.32	0.50	0.14	0.36
3	*0.89*	1.00	0.78	*0.22*	**28**	0.57	0.85	0.28	0.57
4	0.50	0.71	0.28	0.43	**29**	0.35	0.64	0.07	0.57
5	0.67	0.92	0.42	0.50	**30**	*0.92*	1.00	0.85	*0.15*
6	0.57	0.78	0.35	0.43	**31**	0.50	0.64	0.35	0.29
7	0.28	0.50	0.07	0.43	**32**	*0.92*	0.92	0.92	*0.00*
8	0.21	0.42	0.00	0.42	**33**	0.78	0.85	0.71	*0.14*
9	*0.10*	0.21	0.00	*0.21*	**34**	0.25	0.50	0.00	0.50
10	0.64	0.71	0.57	*0.14*	**35**	0.25	0.50	0.00	0.50
11	0.35	0.57	0.14	0.43	**36**	*0.92*	1.00	0.85	*0.15*
12	0.53	0.71	0.35	0.36	**37**	0.39	0.71	0.07	0.64
13	*0.14*	0.28	0.00	*0.28*	**38**	0.35	0.64	0.07	0.57
14	*0.10*	0.21	0.00	*0.21*	**39**	0.67	0.92	0.42	0.50
15	*0.03*	0.07	0.00	*0.07*	**40**	0.21	0.42	0.00	0.42
16	0.67	1.00	0.35	0.65	**41**	0.42	0.71	0.14	0.57
17	*0.96*	1.00	0.92	*0.08*	**42**	0.64	0.64	0.14	0.50
18	0.64	0.92	0.35	0.57	**43**	0.21	0.42	0.00	0.42
19	0.21	0.28	0.14	*0.14*	**44**	0.39	0.42	0.35	*0.07*
20	0.25	0.50	0.00	0.50	**45**	0.50	0.64	0.35	0.29
21	*0.00*	0.00	0.00	*0.00*	**46**	*0.07*	0.00	0.14	*-0.14*
22	0.21	0.42	0.00	0.42	**47**	0.25	0.35	0.14	*0.21*
23	*0.03*	0.07	0.00	*0.07*	**48**	0.32	0.57	0.07	0.50
24	0.53	0.71	0.35	0.36	**49**	0.25	0.50	0.00	0.50
25	0.46	0.78	0.14	0.64	**50**	*0.89*	0.92	0.85	*0.07*

*Item facility should range from 0.20-0.80.
** Item discrimination should be equal to 0.30 or greater.

Copyright by Mahidol University

Item analysis of the MC-Test (part II)

Item Analysis of the MC-Test									
Item	**IF Total**	**IF Upper**	**IF Lower**	**ID**	**Item**	**IF Total**	**IF Upper**	**IF Lower**	**ID**
51	*1.00*	1.00	1.00	*0.00*	**76**	0.50	0.71	0.28	0.43
52	0.64	0.78	0.50	*0.28*	**77**	0.75	0.78	0.71	*0.07*
53	*1.00*	1.00	1.00	*0.00*	**78**	*0.85*	1.00	0.71	*0.29*
54	0.35	0.64	0.07	0.57	**79**	0.57	0.71	0.42	*0.29*
55	*0.96*	1.00	0.92	*0.08*	**80**	*0.10*	0.21	0.00	*0.21*
56	0.28	0.57	0.00	0.57	**81**	0.57	0.64	0.50	*0.14*
57	0.64	0.85	0.43	0.42	**82**	*0.03*	0.07	0.00	*0.07*
58	0.39	0.64	0.14	0.50	**83**	*0.17*	0.28	0.07	*0.21*
59	0.64	0.64	0.64	*0.00*	**84**	0.35	0.57	0.14	0.43
60	0.42	0.57	0.28	*0.29*	**85**	*0.96*	1.00	0.92	*0.08*
61	0.21	0.42	0.00	0.42	**86**	*0.92*	1.00	0.85	*0.15*
62	0.60	0.64	0.57	*0.07*	**87**	0.25	0.50	0.00	0.50
63	*1.00*	1.00	1.00	*0.00*	**88**	*0.10*	0.21	0.00	*0.21*
64	0.78	0.78	0.78	*0.00*	**89**	*0.82*	0.92	0.71	*0.21*
65	0.67	0.64	0.71	*-0.07*	**90**	*0.89*	0.92	0.85	*0.07*
66	0.75	0.78	0.71	0.07	**91**	*0.82*	0.85	0.78	*0.07*
67	0.21	0.42	0.00	0.42	**92**	0.35	0.50	0.21	*0.29*
68	*0.07*	0.14	0.00	*0.14*	**93**	*0.92*	1.00	0.85	*0.15*
69	0.35	0.64	0.07	0.57	**94**	0.42	0.78	0.07	0.71
70	0.53	0.92	0.14	0.78	**95**	0.35	0.57	0.14	0.43
71	*0.82*	0.92	0.71	*0.21*	**96**	0.50	0.57	0.42	*0.15*
72	0.71	0.85	0.57	*0.28*	**97**	0.28	0.57	0.00	0.57
73	0.64	1.00	0.23	0.77	**98**	0.28	0.50	.07	0.43
74	0.46	0.64	0.28	0.36	**99**	*0.03*	0.07	0	*0.07*
75	0.60	0.85	0.35	0.50	**100**	*0.89*	0.92	.07	*0.85*

*Item facility should range from 0.20-0.80.
** Item discrimination should be equal to 0.30 or greater.

Copyright by Mahidol University

Item analysis of the NMC-Test (part I)

Item Analysis of the NMC-Test									
Item	**IF Total**	**IF Upper**	**IF Lower**	**ID**	**Item**	**IF Total**	**IF Upper**	**IF Lower**	**ID**
1	0.60	0.85	0.35	0.50	**26**	0.78	0.85	0.71	*0.14*
2	*0.92*	1.00	0.85	*0.15*	**27**	0.78	0.85	0.71	*0.14*
3	0.14	0.28	0.00	*0.28*	**28**	0.46	0.71	0.21	0.50
4	0.64	0.85	0.42	0.43	**29**	0.60	0.78	0.42	0.36
5	0.21	0.28	0.21	*0.07*	**30**	0.32	0.50	0.14	0.36
6	0.35	0.71	0.00	0.71	**31**	0.21	0.42	0.00	0.42
7	0.67	1.00	0.35	0.65	**32**	0.25	0.50	0.00	0.50
8	0.60	0.92	0.28	0.64	**33**	0.39	0.64	0.14	0.50
9	0.80	1.00	0.57	0.43	**34**	0.21	0.42	0.00	0.42
10	0.28	0.50	0.07	0.43	**35**	0.32	0.64	0.00	0.64
11	0.21	0.42	0.00	0.42	**36**	0.32	0.57	0.07	0.50
12	0.64	1.00	0.28	0.72	**37**	0.25	0.50	0.00	0.50
13	0.32	0.50	0.14	0.36	**38**	0.71	1.00	0.42	0.58
14	*0.03*	0.07	0.00	*0.07*	**39**	0.32	0.64	0.00	0.64
15	0.78	1.00	0.57	0.43	**40**	*0.92*	1.00	0.85	*0.15*
16	0.71	0.85	0.57	*0.28*	**41**	0.39	0.41	0.07	0.34
17	0.21	0.42	0.00	0.42	**42**	0.57	0.85	0.28	0.57
18	0.64	0.92	0.35	0.57	**43**	0.39	0.71	0.07	0.64
19	0.42	0.71	0.14	0.57	**44**	0.71	0.92	0.50	0.42
20	*0.92*	1.00	0.85	*0.15*	**45**	*0.92*	1.00	0.85	*0.15*
21	*0.00*	0.00	0.00	*0.00*	**46**	*0.85*	0.92	0.78	*0.14*
22	0.64	0.85	0.42	0.43	**47**	*0.85*	0.92	0.78	*0.14*
23	*0.00*	0.00	0.00	*0.00*	**48**	0.78	0.92	0.64	*0.28*
24	*0.89*	1.00	0.78	*0.22*	**49**	0.42	0.64	0.21	0.43
25	0.42	0.71	0.14	0.57	**50**	0.60	0.92	0.28	0.64

*Item facility should range from 0.20-0.80.
** Item discrimination should be equal to 0.30 or greater.

Copyright by Mahidol University

Item analysis of the NMC-Test (part II)

Item Analysis of the NMC-Test									
Item	**IF Total**	**IF Upper**	**IF Lower**	**ID**	**Item**	**IF Total**	**IF Upper**	**IF Lower**	**ID**
51	*0.96*	1.00	0.92	*0.08*	**76**	0.42	0.64	0.21	0.43
52	*1.00*	1.00	1.00	*0.00*	**77**	*0.85*	0.92	0.78	*0.14*
53	*0.89*	1.00	0.78	*0.22*	**78**	0.78	1.00	0.57	0.43
54	0.52	0.92	0.00	0.92	**79**	0.67	0.78	0.57	*0.21*
55	*0.92*	0.92	0.71	*0.21*	**80**	0.21	0.35	0.07	*0.28*
56	0.67	1.00	0.35	0.65	**81**	0.53	0.85	0.21	0.64
57	0.60	1.00	0.21	0.79	**82**	0.50	0.71	0.28	0.43
58	0.75	1.00	0.50	0.50	**83**	0.32	0.50	0.14	0.36
59	0.67	0.92	0.42	0.50	**84**	*0.85*	1.00	0.71	*0.29*
60	0.75	0.92	0.57	0.35	**85**	*0.78*	0.85	0.71	*0.14*
61	0.57	0.92	0.21	0.71	**86**	0.42	0.57	0.28	0.29
62	0.21	0.42	0.00	0.42	**87**	*0.85*	0.85	0.85	*0.00*
63	*0.96*	1.00	0.92	*0.08*	**88**	0.52	0.85	0.07	0.78
64	*0.89*	1.00	0.78	*0.22*	**89**	*0.85*	0.85	0.85	*0.00*
65	*0.85*	0.92	0.78	*0.14*	**90**	0.60	0.92	0.28	0.64
66	0.60	0.92	0.28	0.64	**91**	*0.07*	0.14	0.00	*0.14*
67	0.78	1.00	0.57	0.43	**92**	*0.14*	0.21	0.07	*0.14*
68	0.75	1.00	0.50	0.50	**93**	*0.96*	1.00	0.92	*0.08*
69	0.71	0.92	0.50	0.42	**94**	0.52	0.64	0.28	0.36
70	0.39	0.71	0.07	0.64	**95**	*0.03*	0.07	0.00	*0.07*
71	*0.90*	1.00	0.85	*0.15*	**96**	0.21	0.42	0.00	0.42
72	*0.96*	1.00	0.92	*0.08*	**97**	*0.78*	0.78	0.78	*0.00*
73	*0.89*	1.00	0.78	*0.22*	**98**	0.21	0.42	0.00	0.42
74	*0.89*	0.92	0.85	*0.07*	**99**	*0.14*	0.28	0.00	*0.28*
75	*0.85*	1.00	0.71	*0.29*	**100**	*0.03*	0.07	0.00	*0.07*

*Item facility should range from 0.20-0.80.
** Item discrimination should be equal to 0.30 or greater.

Copyright by Mahidol University

Item analysis of the C-Test (part I)

Item Analysis of the C-Test									
Item	**IF Total**	**IF Upper**	**IF Lower**	**ID**	**Item**	**IF Total**	**IF Upper**	**IF Lower**	**ID**
1	*1.00*	1.00	1.00	*0.00*	**26**	*0.96*	1.00	0.92	*0.08*
2	0.64	1.00	0.28	0.72	**27**	0.21	0.35	0.07	*0.28*
3	*0.92*	1.00	0.85	*0.15*	**28**	0.42	0.78	0.07	0.71
4	*0.82*	0.92	0.71	*0.21*	**29**	*0.82*	0.85	0.78	*0.07*
5	0.75	1.00	0.50	0.50	**30**	*1.00*	1.00	1.00	*0.00*
6	0.78	1.00	0.57	0.43	**31**	0.57	0.85	0.28	0.57
7	0.46	0.92	0.00	0.92	**32**	0.50	0.78	0.21	0.57
8	0.35	0.71	0.00	0.71	**33**	*0.89*	0.92	0.85	*0.07*
9	0.42	0.85	0.00	0.85	**34**	0.71	1.00	0.42	0.58
10	0.39	0.64	0.14	0.50	**35**	0.50	0.57	0.42	*0.15*
11	0.46	0.71	0.21	0.50	**36**	0.78	0.92	0.64	*0.28*
12	*0.85*	1.00	0.71	*0.29*	**37**	*0.96*	1.00	0.92	*0.08*
13	0.35	0.71	0.00	0.71	**38**	0.53	0.92	0.14	0.78
14	0.39	0.71	0.07	0.64	**39**	*0.85*	1.00	0.71	*0.29*
15	0.39	0.71	0.07	0.64	**40**	0.42	0.85	0.00	0.85
16	0.67	1.00	0.35	0.65	**41**	*0.85*	0.92	0.78	*0.14*
17	0.78	1.00	0.57	0.43	**42**	0.75	1.00	0.56	0.44
18	0.57	0.85	0.28	0.57	**43**	0.25	0.50	0.00	0.50
19	0.50	0.85	0.14	0.71	**44**	0.21	0.42	0.00	0.42
20	0.28	0.57	0.00	0.57	**45**	0.25	0.50	0.00	0.50
21	0.25	0.50	0.00	0.50	**46**	0.46	0.71	0.21	0.50
22	0.39	0.71	0.07	0.64	**47**	0.42	0.85	0.00	0.85
23	0.21	0.42	0.00	0.42	**48**	0.71	1.00	0.42	0.58
24	0.64	1.00	0.28	0.72	**49**	0.50	0.92	0.07	0.85
25	0.50	0.57	0.42	*0.15*	**50**	*0.92*	1.00	0.85	*0.15*

*Item facility should range from 0.20-0.80.
** Item discrimination should be equal to 0.30 or greater.

Copyright by Mahidol University

Item analysis of the C-Test (part II)

Item Analysis of the C-Test									
Item	**IF Total**	**IF Upper**	**IF Lower**	**ID**	**Item**	**IF Total**	**IF Upper**	**IF Lower**	**ID**
51	*1.00*	1.00	1.00	*0.00*	**76**	0.64	0.85	0.42	0.43
52	0.64	0.85	0.42	0.43	**77**	*0.96*	1.00	0.92	*0.08*
53	*1.00*	1.00	1.00	*0.00*	**78**	0.71	1.00	0.42	0.58
54	0.32	0.64	0.00	0.64	**79**	*0.92*	1.00	0.85	*0.15*
55	*1.00*	1.00	1.00	*0.00*	**80**	0.28	0.57	0.00	0.57
56	0.42	0.85	0.00	0.85	**81**	*0.96*	1.00	0.92	*0.08*
57	0.50	0.78	0.21	0.57	**82**	*0.03*	0.07	0.00	*0.07*
58	0.39	0.71	0.07	0.64	**83**	0.46	0.92	0.00	0.92
59	0.67	0.78	0.05	0.73	**84**	0.28	0.57	0.00	0.57
60	0.50	0.78	0.21	0.57	**85**	0.35	0.78	0.64	*0.14*
61	0.78	1.00	0.57	0.43	**86**	0.67	0.85	0.50	0.35
62	0.42	0.85	0.00	0.85	**87**	0.50	1.00	0.00	1.00
63	*0.96*	0.92	1.00	*-0.08*	**88**	0.28	0.50	0.00	0.50
64	*0.96*	1.00	0.92	*0.08*	**89**	0.75	0.92	0.57	0.37
65	0.75	1.00	0.50	0.50	**90**	*0.82*	0.92	0.71	*0.21*
66	0.60	0.85	0.35	0.50	**91**	*0.85*	1.00	0.71	*0.29*
67	*0.85*	1.00	0.71	*0.29*	**92**	*0.96*	1.00	0.92	*0.08*
68	0.42	0.78	0.07	0.71	**93**	0.21	0.42	0.00	0.42
69	0.78	0.92	0.64	*0.28*	**94**	0.46	0.92	0.00	0.92
70	0.64	0.92	0.35	0.57	**95**	0.53	0.85	0.21	0.64
71	*0.85*	1.00	0.71	*0.29*	**96**	0.60	0.78	0.42	0.36
72	0.75	1.00	0.50	0.50	**97**	*0.89*	0.85	0.92	*-0.07*
73	0.50	0.85	0.14	0.71	**98**	*0.92*	1.00	0.85	*0.15*
74	0.67	0.92	0.42	0.50	**99**	*0.85*	1.00	0.71	*0.29*
75	0.60	1.00	0.21	0.79	**100**	*0.85*	1.00	0.71	*0.29*

*Item facility should range from 0.20-0.80.
** Item discrimination should be equal to 0.30 or greater

Copyright by Mahidol University

Item analysis of the NC-Test (part I)

Item Analysis of the NC-Test									
Item	**IF Total**	**IF Upper**	**IF Lower**	**ID**	**Item**	**IF Total**	**IF Upper**	**IF Lower**	**ID**
1	0.42	0.69	0.15	0.54	**26**	0.53	0.69	0.38	0.31
2	0.80	1.00	0.61	0.39	**27**	0.69	1.00	0.23	0.77
3	*0.00*	0.00	0.00	*0.00*	**28**	0.73	0.92	0.53	0.39
4	*0.88*	0.92	0.84	*0.08*	**29**	0.57	0.61	0.53	*0.08*
5	0.46	0.76	0.15	0.61	**30**	*0.00*	0.00	0.00	*0.00*
6	0.23	0.46	0.00	0.46	**31**	0.61	0.84	0.38	0.46
7	0.80	1.00	0.61	0.39	**32**	0.38	0.46	0.30	*0.16*
8	*0.96*	1.00	0.92	*0.08*	**33**	*0.96*	1.00	0.92	*0.08*
9	0.73	1.00	0.46	0.54	**34**	0.23	0.46	0.00	0.46
10	0.42	0.76	0.07	0.69	**35**	0.50	0.84	0.15	0.69
11	0.34	0.61	0.07	0.54	**36**	0.57	0.84	0.30	0.54
12	0.61	1.00	0.23	0.77	**37**	0.38	0.43	0.07	0.36
13	0.50	0.84	0.15	0.69	**38**	0.42	0.69	0.15	0.54
14	*0.03*	0.07	0.00	*0.07*	**39**	0.73	0.84	0.61	*0.23*
15	0.30	0.53	0.07	0.46	**40**	*0.92*	1.00	0.84	*0.16*
16	0.73	0.84	0.76	*0.06*	**41**	0.57	0.92	0.23	0.69
17	0.45	0.84	0.00	0.84	**42**	0.46	0.46	0.46	*0.00*
18	*0.96*	1.00	0.92	*0.08*	**43**	0.57	0.76	0.38	0.38
19	0.23	0.46	0.00	0.46	**44**	0.46	0.76	0.15	0.61
20	0.65	0.76	0.53	*0.23*	**45**	*0.92*	1.00	0.84	*0.16*
21	0.23	0.46	0.00	0.46	**46**	0.80	1.00	0.76	*0.24*
22	0.73	0.92	0.84	*0.08*	**47**	0.73	1.00	0.46	0.54
23	0.23	0.38	0.07	0.31	**48**	0.57	1.00	0.15	0.85
24	*0.92*	1.00	0.84	*0.16*	**49**	0.61	0.69	0.53	*0.16*
25	*0.80*	0.84	0.76	*0.08*	**50**	0.34	0.38	0.30	*0.08*

*Item facility should range from 0.20-0.80.
** Item discrimination should be equal to 0.30 or greater.

Copyright by Mahidol University

Item analysis of the NC-Test (part II)

Item Analysis of the NC-Test									
Item	**IF Total**	**IF Upper**	**IF Lower**	**ID**	**Item**	**IF Total**	**IF Upper**	**IF Lower**	**ID**
51	*1.00*	1.00	1.00	*0.00*	**76**	0.57	0.84	0.30	0.54
52	*0.96*	0.92	1.00	*-0.08*	**77**	0.65	1.00	0.30	0.70
53	*0.92*	0.92	0.92	*0.00*	**78**	*0.88*	0.92	0.84	*0.08*
54	0.34	0.61	0.07	0.54	**79**	*1.00*	1.00	1.00	*0.00*
55	0.80	1.00	0.61	0.39	**80**	0.76	0.92	0.61	0.31
56	0.38	0.61	0.15	0.46	**81**	0.25	0.30	0.00	0.30
57	0.26	0.53	0.00	0.53	**82**	0.25	0.76	0.15	0.61
58	0.26	0.53	0.00	0.53	**83**	0.42	0.76	0.00	0.76
59	0.42	0.69	0.15	0.54	**84**	0.61	1.00	0.23	0.77
60	0.53	1.00	0.07	0.93	**85**	*1.00*	1.00	1.00	*0.00*
61	0.61	0.84	0.30	0.54	**86**	0.38	0.61	0.15	0.46
62	0.30	0.61	0.00	0.61	**87**	0.23	0.46	0.00	0.46
63	0.76	0.84	0.69	*0.15*	**88**	0.76	1.00	0.53	0.47
64	*0.84*	0.92	0.76	*0.16*	**89**	0.61	0.76	0.46	0.30
65	*0.96*	1.00	0.92	*0.08*	**90**	*0.88*	1.00	0.76	*0.24*
66	0.50	0.92	0.07	0.85	**91**	0.80	0.92	0.76	*0.16*
67	0.80	1.00	0.69	0.31	**92**	0.26	0.53	0.00	0.53
68	0.69	1.00	0.38	0.62	**93**	*0.96*	1.00	0.92	*0.08*
69	0.80	1.00	0.69	0.31	**94**	*0.88*	0.92	0.84	*0.08*
70	0.80	1.00	0.69	0.31	**95**	0.30	0.46	0.07	0.39
71	0.80	0.92	0.76	*0.16*	**96**	0.34	0.69	0.00	0.69
72	0.73	0.84	0.61	*0.23*	**97**	0.80	1.00	0.61	0.39
73	*1.00*	1.00	1.00	*0.00*	**98**	0.38	0.76	0.00	0.76
74	*0.88*	1.00	0.76	*0.24*	**99**	0.61	0.84	0.38	0.46
75	*1.00*	1.00	1.00	*0.00*	**100**	0.23	0.46	0.00	0.46

*Item facility should range from 0.20-0.80.
** Item discrimination should be equal to 0.30 or greater.

Copyright by Mahidol University

APPENDIX I

TRANSCRIPT OF HIGH-LANGUAGE-ABILITY GROUP

Researcher: Good afternoon, how are you? My name is Montarat Rungruangthum. But you can call me P'Orm. I'm a graduate student in Applied Linguistics at the Faculty of Science, Mahidol University. My research study focuses on the four cloze tests. So I'd like to have a group interview with you to study what test-taking strategies you used while taking each of the four cloze tests. Take a look at this type of the cloze test from one of your cloze test papers. It's called the C-test in which the *second part* of every *second word* is deleted. The next type is the NC-Test, you will see that the *second part* of every *third word* is removed. And this is the MC-Test. I deleted the *first part* of every *second word* in this cloze test. And this one is the MC-Test. The *first part* of every *third word* is deleted. Do you have any questions? [*Pause*] NO. OK? This group interview will take about 30 to 40 minutes. Would you allow me to record your answer during the interview?

C-Test-taker: Yes.

NMC-Test-taker: No Problem

NC-Test-taker: OK

MC-Test-taker: Um…. Yes.

Researcher: Thanks a lot. Please take a look at this interview format. Then check what type of cloze test you took: the C-Test, the NC-Test, the MC-Test, or the NMC-Test. Write it down on your interview format. [*Pause*] Good. Now who's gonna be the first one to answer my interview question? [*The NMC-Test-taker made a sign.*] OK. My question is "what strategies did you used while taking the NMC-Test?"

NMC-Test-taker: I read from the beginning to the end to find the keywords in each passage. I felt stressed because I had never taken such a test before. What a test! It gave only the second part of the deleted words.

Copyright by Mahidol University

Researcher: So could you give an example how you could restore each test item?

NMC-Test-taker: I read the whole passage before restoring the deleted part. I had some prior knowledge about some passages in the test so I felt familiar to the content of the test. I also analyzed part of speech of each keyword in each sentence.

Researcher: Did you use any other strategies?

NMC-Test-taker: For the third and the fourth passages, I read most of the sentences in part because the test contents seemed to be general knowledge.

Researcher: How about the second passage?

NMC-Test-taker: I read through the passage. I knew that the text was about computer. But I did not know so many words that I had to use the contextual clues to find the meanings of these words. Moreover, for some items, I studied the functions of some deleted words such as logical connectors.

Researcher: Would you like to add anything else?

NMC-Test-taker: No.

Researcher: Thank you. [*Turning to the C-Test-taker*] And how about you? What strategies did you use while taking the C-Test?

C-Test-taker: I translated a few sentences at the beginning and at the end of the paragraph into Thai to find the main idea. In this way, it helped me understand the content much better.

Researcher: Did you use any other strategies to fill in the deleted parts?

C-Test-taker: For the third passage, the content is smoking. I used the contextual clue and the grammatical structure to check whether the passage referred to smoking [*Pointing to item 58*] or smoker [*Pointing to item 59*].

Researcher: What about the other passages?

C-Test-taker: I read for the main ideas, and I also used contextual clues in each sentence.

Researcher: How? Could you give me some examples?

Copyright by Mahidol University

C-Test-taker: In the fourth passage, I comprehend this sentence to find the answer for this item [*Pointing to item 80*]. At first, I though it would be **popul a t i o n** but when I read it more carefully, I then knew that it should be "**popul a r i t y**"

Researcher: Thank you. [*Turning to the MC-Test-taker*] How about you? What strategies did you used while taking the MC-Test?

MC-Test-taker: I read every sentence from the beginning to the end of the stories before filling in each blank. I thought I could find more correct answers for the deleted items.

Researcher: Did you use any other strategies?

MC-Test-taker: The contents of some passages were so familiar to me that I could answer those items. I also translated some words into Thai. But some items, such as this one [*Pointing to item 17*] with one blank space and one letter, I knew immediately that it must be "**o f**". For pronouns and prepositions, I could also restore immediately. But sometimes, I had to analyze the deleted words to find which type of word is missing.

Researcher: [*Turning to the NC-Test-taker*] What strategies did you use to fill in the NC-Test?

NC-Test-taker: Mostly, I focused on the meanings of the deleted words.

Researcher: How?

NC-Test-taker: I often used grammatical structures and part of speech, and translated each sentence with the missing words.

Researcher: Did you read from the beginning to the end, or just in part, or jump around?

NC-Test-taker: I read only two or three sentences at the beginning of the paragraph, trying to find the main idea.

Researcher: [*Turning to the C-Test-taker*] Did you read from the beginning to the end, or only parts of the test, or jump around?

C-Test-taker: For myself, I read only some parts of the test because I was afraid that I could not finish the test on time.

Researcher: [*Turning to the MC-Test-taker*] Did you read from the beginning to the end, or only parts of the test, or jump around?

Copyright by Mahidol University

MC-Test-taker: I read line-by-line as I mentioned earlier or read some parts of the test.

Researcher: [*Turning to the NMC-Test-taker]* How about you?

NMC-Test-taker: I read the whole passage to understand what it was about. I chose this strategy because this test was difficult. I also read only some sentences in Passages three and four in order to find part of speech of each deleted item.

Researcher: Did you use any other strategies to figure out the answer in the NMC-Test?

NMC-Test-taker: Uh…. I'm afraid that my answer might not be correct so I jotted down the word I guessed in pencil for double-check later. Anyway, if I did not know the answer on any item, I would leave it blank.

Researcher: [*Turning to the MC-Test-taker*] And you?

MC-Test-taker: I couldn't restore about 15 to 16 items because I couldn't figure out. In addition, some answers I guessed were wrong.

Researcher: Did you guys use any other ways to find the answer in the cloze tests?

C-Test-taker: Well, if the first part of each item is deleted, it is easy to guess. I had a lot of fun playing such a Puzzle game. However, if I didn't know the answers, I would leave them blank.

Researcher: And what about your strategy?

NC-Test-taker: First, I counted the number of the deleted letters on each item. Then, I guessed. But if I couldn't restore any items, I would ignore it.

Researcher: Finally, the interview's finished. Your answers are very helpful to my research. Thank you so much again for your kind participation. I appreciate it a lot.

NOTE: Transcribed as recorded; not edited to correct language errors

Copyright by Mahidol University

TAPESCRIPT OF LOW-LANGUAGE-ABILITY GROUP

Researcher: Good afternoon, Folks. My name is Montarat Rungruangthum or P' Orm. I'm a graduate student in a Master's of Arts in Applied Linguistics at the Faculty of Science, Mahidol University. My research focuses on the cloze tests and the test-taking strategies of these four cloze tests. OK. Let's start now. There are your cloze test papers and the interview formats. What I'd like to learn from you is about your strategies while you were taking each of the four cloze tests. Please fill in the interview formats and then check what type of the cloze test you took: the C-Test, the NC-Test, the MC-Test, or the NMC-Test. You can see the test type mentioned on the front page of your test paper. Before the interview starts, let me explain each type of the cloze test in short. So you will understand each type of the cloze test. OK. Let's see the first type of the cloze test. It is called the C-Test. Anyone who's got the C-Test paper, please show it to your friends. The C-Test is an original cloze test which the *second part* of every *second word* is deleted. The one who got the NC-Test, please show it to your friend. The second type is the NC-Test is a test in which the *second part* of every *third word* is removed. The third type is the MC-Test that the *first part* of every *second word* is deleted. May we look at your test paper of the NMC-Test? The fourth type is the NMC-Test. I deleted the first part of every third word from the passage. Any questions? [*Pause*]. For the group interview, it will take around 30 to 40 minutest before the interview starts. First of all, would you allow me to record your responses during the interview?

C-Test-taker: No problem.

NC-Test-taker: You can do it.

MC-Test-taker: Go ahead.

NMC-Test-taker: Sure

Copyright by Mahidol University

Researcher: Now, please take a look at your test paper and I allow you five minutes to recall what strategies you used while taking the test. And then, I d' like you to explain your test-taking strategies to me one by one. Besides tape-recording, I will take not your interview responses as well. Who wants to be the first one to answer my questions?

Researcher: [*Turning to the first interviewee*] Well, what strategies did you use while taking the deleted parts in the NC-Test

NC-Test-taker: Well, I used the clues preceding each item and also figure out what part of speech of each item should be. And then I tried to translate the meaning of each deleted word into Thai. For some items, I could restore them immediately by just looking at the deleted words only because these deleted words were articles and prepositions.

Researcher: Could you give me a few examples?

NC-Test-taker: [*Pointing to item 24*] Look at this item. The answer is "a" because there is only one blank space for this word. So it must be an article "a".

Researcher: And did you use any other strategies?

NC-Test-taker: I also used the context clues in the sentence

NC-Test-taker: Now, take a look at item 21. I saw the word "as" following the deleted word. I know that the word "as" can collocated with "such". The fist part of this item is "su". So the correct word must be "**suc h**". Mostly, I used translation into Thai to find the meaning of each word.

Researcher: Anything else?

NC-Test-taker: Uh… Just analyzing parts of speech of the deleted words as mentioned earlier.

Researcher: [*Turning to the third interviewee*] How about you? What strategies did you use while you were taking this MC-Test?

MC-Test-taker: Well, while I was taking the MC-Test, I would not figure out which word I should fill in. It was quite difficult. I read each sentence in the paragraph to see what it was about. For example, the first passage was so difficult that I could answer only a few items.

Copyright by Mahidol University

Researcher: Which item did you answer correctly? And how could you restore it?

MC-Test-taker: [*Pointing to item 6*] here**.** I used the clues preceding the deleted words. I saw the word "example" so I know that the word "**f or**" could collocate with "example". For item 19, the part of speech of this item must be a verb because the verb of the sentence is missing. It maybe "has" or "was", but the tense of this passage used present simple so the correct answer for this item should be "has".

Researcher: Did you use any other strategies that you used while taking the MC-Test?

MC-Test-taker: I was looking at the first two or three sentences of each passage. In passage three, I understood that it was about smoking so I could use my previous knowledge of vocabulary to help me fill in the deleted part correctly. So my answer for item 56 was "**s m o king**".

Researcher: [*Asking the fourth interviewee*] And you?

NMC-Test-taker: The same.

Researcher: Could you explain how you could find the answer for the deleted words in the NMC-Test?

NMC-Test-taker: Like the strategies of the interviewee taking the NC-Test. I looked at the deleted word first. And then, I could think of the word to fill in the blank, such as in item 38, I guess that the answer should be "**t h eir**". But for many items, I could not guess by just looking at the deleted word. I had to use the contextual clues as well as grammatical structure of each sentence to help me think of a suitable word for each item.

Researcher: Could you please give me an example?

NMC-Test-taker: OK. For item 46, I saw only the letter "d" at the end of the deleted word so I could not figure it out. I tried to study the part of speech of this item. I assumed that the deleted word must be a conjunction so the answer should be "**a nd**". But for item 34, I used contextual clues which are "verb to be' an "in" so I restored it by "**i n t e rested**". I think that filling the item with the fist part deleted was very difficult because it seemed unfamiliar to me.

Copyright by Mahidol University

Researcher: Did you use the other strategies to fill in the deleted words?

NMC-Test-taker: None

Researcher: [*Turning to the first interviewee*] How about you? Which strategies did you used while taking the C-Test?

C-Test-taker: I guessed if I didn't know the answer. It seemed like playing "Hangman". I mean I had to find part of speech and the find the meaning of each item. For example, item 15, I could finally restore "**vir u s e s**" but I shouldn't have misspelled it.

Researcher: Did you read from the beginning to the end, or read some parts of the test, or read jump around while taking the C-Test?

C-Test-taker: If the passage were fairly difficult, I would read a few sentences at the beginning of the paragraph to find the main idea of the passage. In this way, this I would understand what the passage was about.

NMC-Test-taker: Just the same in that I had to read a few sentences at the beginning and at the end of the passage. If the passages were difficult, such as passage one and passage two, I had to read through the passages once. And then, I tried to think of a suitable word for each item.

Researcher: [*Asking the other interviewees*] Did both of you use the same strategies as these two interviewees?

NMC-Test-taker: Not really. Mostly, I read just some parts of sentences in each passage. Except passage one, I had to read the whole passage before filling in the deleted words.

Researcher: [*Turning to the second interviewee*] And you? What strategies did you frequently used?

NC-Test-taker: I partially read the first and the second sentences because the main idea of each sentence was generally presented at the beginning or at the end of the passage.

Researcher: If you could not restore any item, did you guess or find another way to fill in that item?

NC-Test-taker: First, I counted the number of the deleted letters of each item to help me guess. If I could not figure it out, I left it blank.

Copyright by Mahidol University

Researcher: [*Turning to the fourth interviewee*] And you?

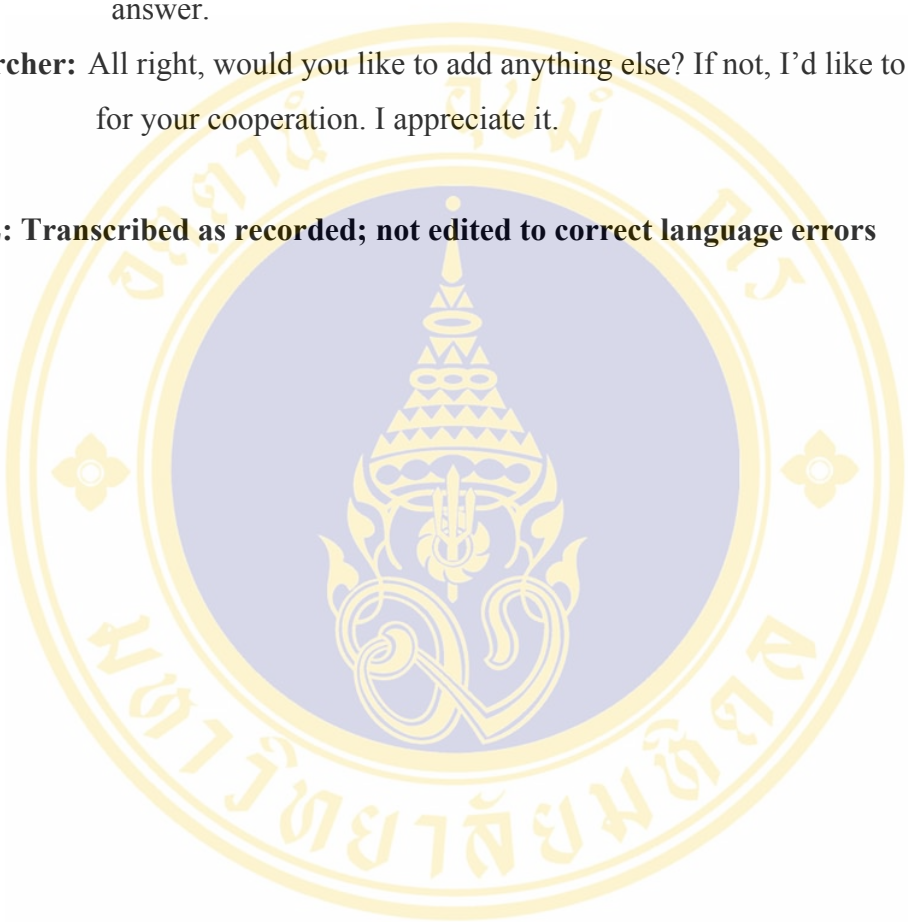
NMC-Test-taker: Just left it blank. I didn't restore at all because I didn't know what the answer was.

MC-Test-taker: I could not guess anymore so I left it blank.

C-Test-taker: I tried to find the similar words used in the other passages to help me the answer.

Researcher: All right, would you like to add anything else? If not, I'd like to thank you for your cooperation. I appreciate it.

NOTE: Transcribed as recorded; not edited to correct language errors



Copyright by Mahidol University

BIOGRAPHY

NAME	Miss Montarat Rungruangthum
DATE OF BIRTH	29 May, 1978
PLACE OF BIRTH	Bangkok, Thailand
INSTITUTIONS ATTENDED	Thammasat University, 1996-2000 Bachelor of Arts (Linguistics) Mahidol University, 2001-2005 Master of Art (Applied Linguistics)
HOME ADDRESS	263 Rama 6 Road Soi 20 Ratchatewee Bangkok, 10400 Tel 04-1463-405, 02-2158-211

Copyright by Mahidol University