

LOCATING COLLEGE STUDENTS' READING GRADE LEVEL THROUGH CRITERION-REFERENCED VOCABULARY TESTS

Tzung-yu Cheng

Associate Professor, The Center for General Education, China Medical
University

Abstract

This study is intended to locate the English reading grade levels of EFL college students by using word criterion-referenced tests. The subjects are 145 college freshmen at a medical university in Taiwan. These students were administered The Slossen Oral Reading Test at mid-term. Descriptive statistics delineates the range to be from Grade 1 to Grade 6. Both Grade 1 and Grade 6 are at the extremities of the bell curve as only 0.7 % of the students were at Grade 1 and 4.8% at Grade 6. The other Grade levels were determined to contain as follows: Grade 2 = 5.5%, Grade 3 = 24.15%, Grade 4 = 38.6% and Grade 5 = 26.2%. The results are followed with a graph to illustrate this range. Discussions are offered to include the advantages and disadvantages of using this approach under the current EFL teaching context in Taiwan. Limitations and further research paths are also suggested.

Keywords: English Reading Tests, English Standardized Tests, Ability Grouping,
Grade Equivalents

Requests for reprints should be sent to Tzung-yu Cheng, The Center for General Education, China Medical University, 91 Hsueh-Shih Road, Taichung 404, Taiwan.
E-mail: tychen@mail.cmu.edu.tw

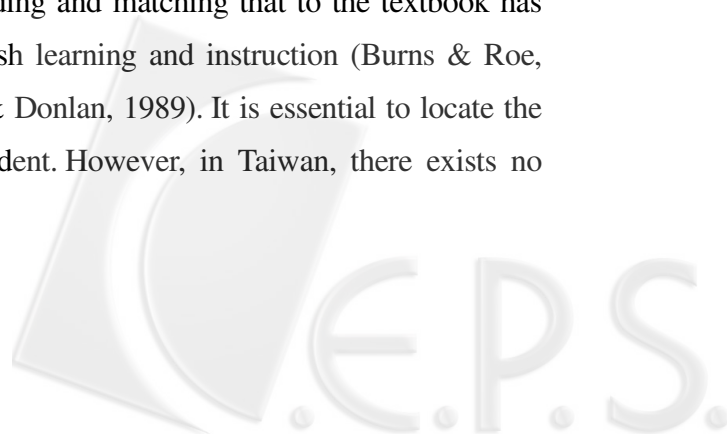


INTRODUCTION

In Taiwan, most college students are required to read content area textbooks in English (Cheng, 1993). According to surveys, 89 % of the college freshmen were required to read content area textbooks originally written for the college students of English native speakers (Tsai, 1978). Further investigation indicated that 94% of the ESL content teachers "use textbooks written in English for either all courses or at least one-third of the courses that they are teaching" (Kong, 1996, p.42). Yet, students experience great difficulty reading those textbooks (Cheng, 1993; Kong, 1996; Tsai, 1978; Yiau, 1993). Cheng (1993) observed a medical college in central Taiwan over a two-year period. He found that medical students do not deal directly with reading textbooks in English; instead, they look for translated versions. Other study also revealed that 85% of the doctoral and graduate students at public universities use repeated reading as a dominating strategy to process the content of academic English (Yiau, 1993).

What caused these reading difficulties? These college texts are written for English speakers who "ideally" should possess at least a reading grade level equivalent to 12th grade. However, there exists no information that determines the grade level English read by the college students in Taiwan. It is a fact that the purpose of learning English is to communicate with the native users of the language and be able to read for understanding and obtaining knowledge explained in English. Therefore, whatever the goals of English instruction in Taiwan from elementary school through high school may be, once the students enter a college, they will face English reading textbooks that are completely the same as those used in U.S. colleges.

Determining the grade level of reading and matching that to the textbook has been proved to facilitate effective English learning and instruction (Burns & Roe, 1989; Lipson & Wixson, 1991; Singer & Donlan, 1989). It is essential to locate the English reading grade level of each student. However, in Taiwan, there exists no



conclusive approach to ascertain an individual student's reading grade level in English. Furthermore, in a typical language classroom, the reading abilities may vary from five to seven grade levels (Singer & Donlan, 1989). As a result, a measure is needed to identify a student's actual reading grade level and select reading material that meets the student's reading ability. This is a challenge to not only the teachers of English as a second language but also to the colleges professing to teach and expand such academic knowledge.

Several popular approaches have been used in the United States to find a student's reading grade level and match it with proper reading materials. Informal Reading Inventory (IRI) (Burns & Roe, 1989), as an example, is one of the commercial informal criterion-referenced reading tests. In an IRI, a representative passage (200 to 500 words, depending on grade level) is selected to represent each of the grade levels from primary to 12th grade. Before giving the comprehensive test, each student is first asked to do a wordlist test to locate the starting grade level. Then the student reads aloud the appropriate passage for a word accuracy score. Next, comprehension questions are asked to calculate a comprehension score. An IRI criteria is then checked to find the test taker's three levels of reading ability: independent, instructional and frustration level (Burns & Roe, 1989). The major problem with this IRI test is that it is individualized, so it is not possible to administer the test to a great number of students at once to save time. In addition, it is common for a student to successfully complete the graded passage at a high level, but fail to complete passages at lower levels. Therefore, it is questionable whether the graded passages can actually represent the intended grade level. Using a single passage to represent each grade level may not be enough as the student might by chance already be familiar with the passage.

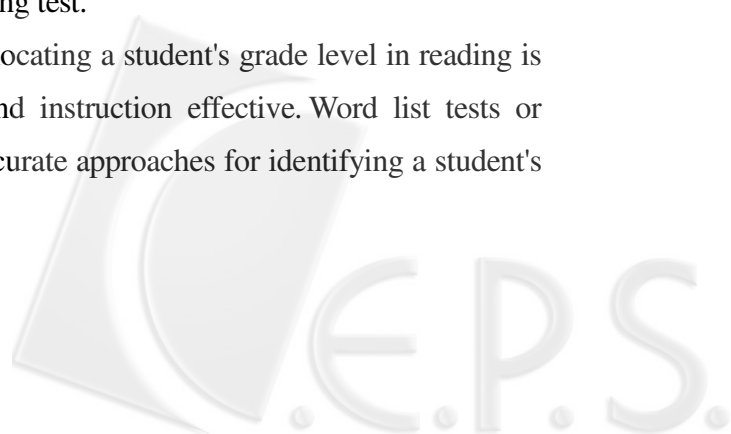
The cloze procedure, an informal reading test, is also a common approach used to find a student's three levels of reading ability: independent, instructional or frustration (Lipson & Wixson, 1991). This test can be administered individually or as a group. Five steps are needed: 1) Selecting three passages above 100 words each; 2)

In each passage keep the first and last sentence intact; 3) Starting with the second sentence, delete every fifth word until there are 20 blanks; 4) Average the scores of three passages; 5) Check the cloze criteria. This test is rather easy to construct; however, with this approach, when new material is to be used in class, the teacher needs to go through the same procedure to make a new test. The three cloze criteria developed using this same procedure vary among each other (Cheng, Sheen & Hung, 1997). This results in some puzzlement on the part of the teacher in choosing which criterion to follow.

The most frequently used approach for locating a student's reading level in the United States is using standardized reading tests such as the Gate-MacGinitie Reading Tests. These tests are developed by following strict statistical procedures and using a population measure to develop a norm. The purpose is to identify whether the student is reading at the grade level or lagging behind or ahead of grade level. The resulting scores may then be used by the teacher to determine what course of study or remedial steps need to be taken to match the student with the corresponding grade level material. A typical standardized test includes two subtests: vocabulary and reading comprehension. This test is easy to use, but having only two forms for each grade level may pose a disadvantage. These tests are also expensive and only authorized personnel can administer and score them.

There is a trend of tests that is less expensive and can be used quickly by a teacher. This is the word list test. They are criterion-referenced tests. The words in each test come from the most frequently used words at each grade level (primary to 12th grade). The effects of using word tests to identify a student's reading level have been well documented. Chall (1987) suggested that word meaning scores are so highly correlated with reading comprehension scores that a reading vocabulary test may be substituted for a paragraph-meaning test.

This research review indicates that locating a student's grade level in reading is essential in making English learning and instruction effective. Word list tests or criterion-referenced tests are easy and accurate approaches for identifying a student's



reading grade level. This study will illustrate the results in Table 1 and Figure 1 with data analysis and further discourse.

METHOD

Participants

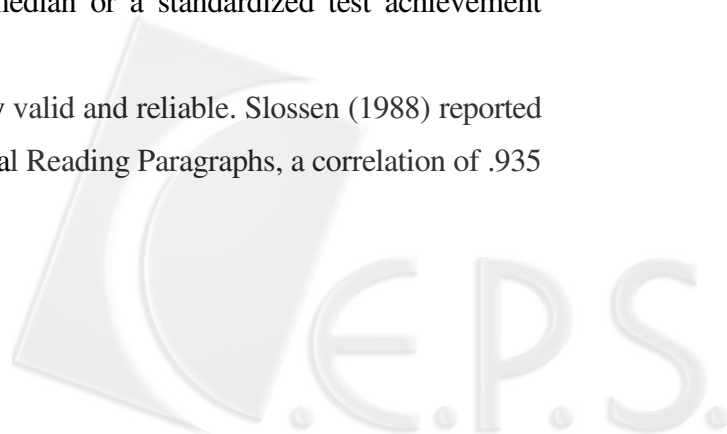
The subjects consisted of 145 freshmen at China Medical University. They were selected from 216 students who enrolled in five English reading classes. These 216 students came from 11 departments. However, for research data to be valid, a minimum number of subjects should be higher than 18 (Gay, 1982). As a result, 145 students were drawn from the pool of the original 216. These students came from five departments: Biotech, Hospital Management, Occupational Safety, Physical Medicine and Nutrition.

This group of subjects does not represent the population of the whole university, as the following departments are not included: Medicine, Dentistry, Pharmacy, Physical Therapy and Chinese Medicine.

Measure

The instrument used in this study is The Slossen Oral Reading Test (SORT) (Slossen, 1988). The test consists of 200 words grouped into 10 levels: P (primer), 1, 2, 3, 4, 5, 6, 7, 8 and 9-12. Each level includes 20 words. In administering the test, a teacher may use it as an individualized test or modify it as a group test by asking the students to read each word aloud for word identification and accuracy or to interpret the meaning of each word for comprehension. For the current study, the students were asked to write down their interpretation of each word. The reading level obtained from this test represents the median or a standardized test achievement level.

The test was considered to be highly valid and reliable. Slossen (1988) reported a correlation of .96 with The Standard Oral Reading Paragraphs, a correlation of .935



with The Peabody Individual Achievement Test of Reading Recognition, and a correlation of .901 with The Woodcock-Johnson Test of Achievement with Letter-Word Identification.

Rating

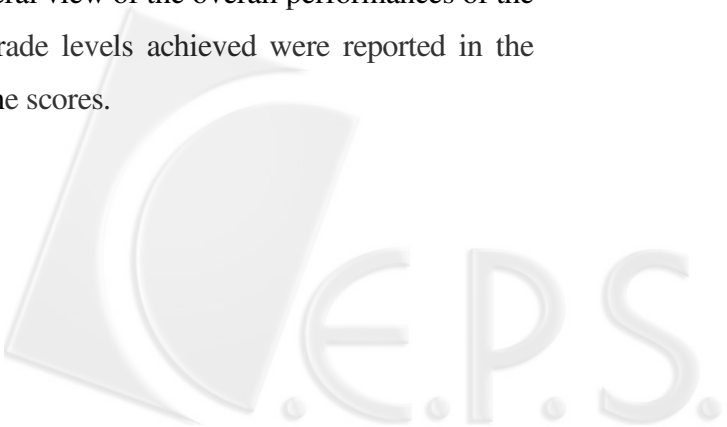
Context is not used as a clue to facilitate comprehension in the test used. In addition, a word may include different meanings; for example, "sound" in physics means "the noise made by nature or people," in geography, it means "a sea bay" or as a verb "to make noise" or as an adjective "perfect" or "sturdy." For this reason, any interpretation touching upon one of the different meanings was counted correct. Two impartial raters graded the interpretations completed by the subjects. Discrepancies in grading were settled by discussion between the raters.

Procedure

This measure was administered as a mid-term test. It was believed that the participants would respond to each of the words more sincerely and carefully. Ninety minutes were given for completing the word list.

Data Analysis

The real purpose of testing is to assist the instruction (Lipson & Wixson, 1991). Therefore, it is not the purpose of the study to compare the difference of the overall achievement of each of the departments. To avoid discrimination by comparison and tagging groups, the study is more interested in finding out individual student's performance in order to explain the student's reading difficulty and to suggest solutions. Therefore, in this study, only descriptive statistics were used to delineate the precise grade level for a general view of the overall performances of the students. The maximum and minimum grade levels achieved were reported in the data along with the percent and mode of the scores.



RESULTS

Distribution of Grade Level Scores

Table 1 presents the frequency and the percent showing the performance of the 145 students tested. Descriptive data indicates that both Grade 1 and Grade 6 are at the extreme ends of the curve, however, there are only 0.7% at grade 1 and 4.8% at Grade 6. A low percentage is also found at Grade 2 with 5.5%. The overall distributions by percent for the remaining grade levels are 26.2% at Grade 5, 38.6% at Grade 4 and 24.15% at Grade 3.

Table 1: Frequency and Percentage

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1.00	1	.7	.7	.7
	2.00	8	5.5	5.5	6.2
	3.00	35	24.1	24.1	30.3
	4.00	56	38.6	38.6	69.0
	5.00	38	26.2	26.2	95.2
	6.00	7	4.8	4.8	100.0
	Total	145	100.0	100.0	

Table 2 lists the median, mode, range, minimum, maximum and percentiles of the results. The median as well as the mode fall at the Grade 4 level. The range is five grade levels from 6 to 1. Grade 6 is at the maximum extreme grade level range and Grade 1 at the minimum extreme grade level range. Grade 5 level is in the upper 75 percentile while grade 3 level shows a performance in the lower 25 percentile.

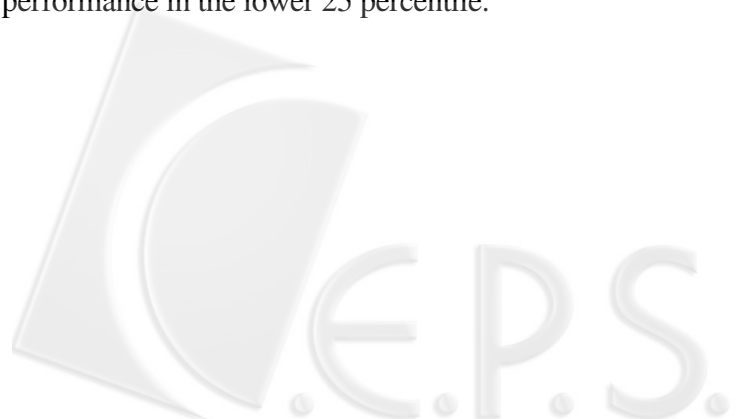


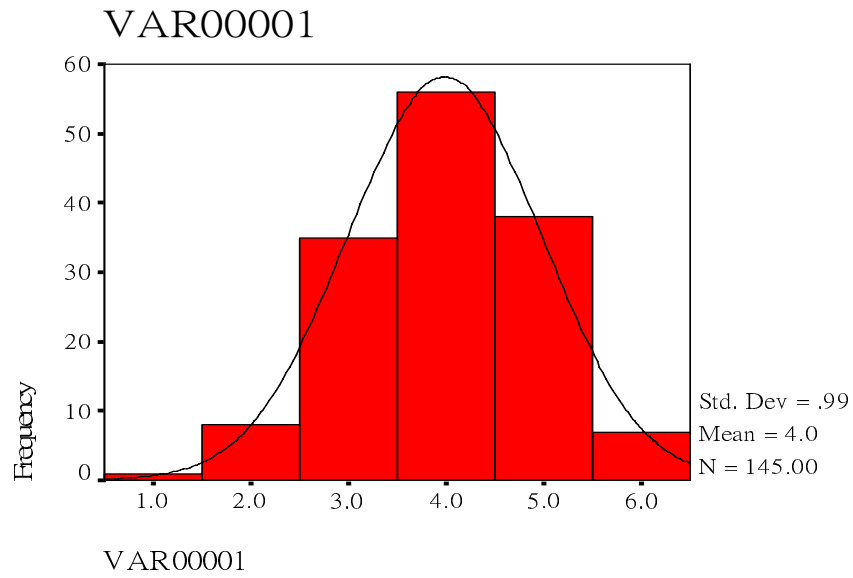
Table 2: Median, Mode, Range, Minimum, Maximum and Percentiles

N	Valid	145
	Missing	0
Median		4.0
Mode		4.0
Range		5.0
Minimum		1.0
Maximum		6.0
Percentiles	25	3.0
	50	4.0
	75	5.0

Figure 1 shows the mean and the normal curve of the overall performance of the 145 participants. The mean is at Grade 4 level. The distribution spreads out in a bell shaped curve showing that the majority of the student subjects fell within grades 3, 4 and 5. A small amount scored at Grade 1 with slightly larger groups in Grade 2 and Grade 6. No participants scored better than Grade 6 level in reading English.



Figure 1: Mean and Normal Curve



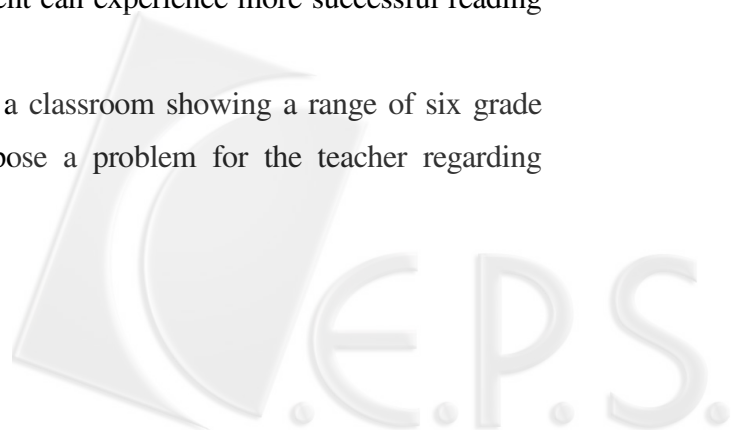
DISCUSSION

Most college students in Taiwan are required to read textbooks written by and for native English speakers. This fact imposes a problem for the college students learning English as a second language as well as their teachers. No study in Taiwan attempts to locate the reading grade level of these students or compare their levels equivalent to those of students in the United States. Therefore, the purpose of this study is to locate college EFL students' reading grade levels. This is measured by administering a criterion-referenced word test. The goals are twofold: 1) to help locate a student's grade level as quickly as possible and 2) to assign reading materials that match the student's reading grade level.

The results of this study reveal that each participant performed below 7th grade level when measured by a criterion-referenced word test. Grades 3, 4 and 5 are shown to be the most common grades of those participating. These scores are equivalent to the elementary grade levels of native speakers of English. Obviously, the participants will meet great difficulties when they attempt to read college content textbooks used in English speaking countries. Findings by Cheng (1993), Kong (1996), Tsai (1978) and Yiau (1993) are consistent with these findings in concluding that students in Taiwan experience great difficulty reading college content English textbooks.

The range of scores among the college student participants shows them spanning five grade levels. The testing has an advantage when using grade levels from k-12 when scoring. This is more precise and covers an extensive range of ability levels. Under this approach, it is rather easy to assign appropriate reading material at the student's reading grade level. Independent reading becomes possible and motivation may increase as the student can experience more successful reading with better comprehension.

The discrepancy of abilities within a classroom showing a range of six grade levels in the reading of English will pose a problem for the teacher regarding



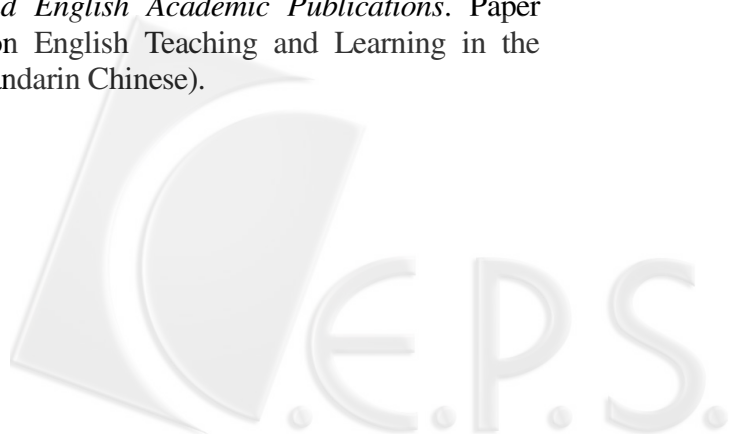
grouping by ability. As in the case of this study, six classes are needed. But the small number of students on both extremes (Grades 1, 2 and 6) certainly need special attention and different reading material. Such a low classroom population with these special students is not cost effective. In addition, the students at Grades 1 and 2 may need further diagnostic teaching and remedial reading material (Lipson & Wixson, 1991). Are there enough teachers who have this specialized training in language diagnostics and remedial education? These problems need to be addressed given the current teaching context in Taiwan.

This study offers an approach for identifying a student's English reading ability. It is fast and easy to use. However, the results are limited to students from five departments who are of similar abilities in English. The college entrance exam English scores usually make this classification of ability. Students from the departments of pharmacy, dentistry, physical therapy, medicine and Chinese medicine may have higher scores. Those departments are usually higher in English ability as their English entrance exam scores would indicate. Students from other higher education institutions may also score at varying grade levels. Such discrepancies may be rectified through the use of a standardized test in which both vocabulary and comprehension in English are measured. The English scores on the college entrance examination can be used to correlate the standardized test. These measures would help teachers place students in English classes with others of their ability and teachers with grade equivalent reading material and textbooks.



REFERENCE

1. Burns, P. C. & Roe, B. D. (1989). *Burns/Roe Informal Reading Inventory – Preprimer to Twelfth Grade*. Boston: Houghton Mifflin.
2. Chall, J. S. (1987). Two Vocabularies for Reading: Recognition and Meaning. In Curtis, M. E., & McKeown, M. G. (Eds.) *The Nature of Vocabulary Acquisition* (pp.7-15). Hillsdale, NJ.: Lawrence Erlbaum Associates.
3. Cheng, T. Y. (1993). The Syntactical Problems Chinese College Students Meet in Reading English Technical Textbooks. *Resource in Education*. (ERIC Document Reproduction Service No. ED 364 094).
4. Cheng, T. Y., Sheen, L. B., & Hung, T. Y. (1996): The Effects of Cloze Reading on EFL Readers' Recall, Inference, and Main Idea Recognizing.第十三屆中華民國英語文教學研討會論文集(pp.145-153)，台北:文鶴出版社。
5. Gay, L. R. (1989). *Educational Research: Competence for Analysis & Application*. Columbus, OH: Merrill Publishing.
6. Kong, H. C. (1996). Teaching College English: From ESL to ESP. *Studies in English and Literature*, 1, 40-45.
7. Lipson, M. Y. & Wixson, K. K. (1991). *Assessment & Instruction of Reading Disability: An Interactive Approach*. New York: HarperCollins Publishers.
8. Singer, H., & Donlan, D. (1989). *Reading and Learning from Texts*. Hilldale, NJ: Lawrence Ehbaum.
9. Slossen, R. L. (1988). *Slossen Oral Reading Test (SORT)*. New York: Slossen Educational Publications.
10. Tsai, C. Y. (1978). *A Study of Freshmen English Teaching in the Republic of China*. (Unpublished MA thesis; National Taiwan Normal University). Taipei: NTNU.
11. Yiau, C. K. (1993). *The Approaches Chinese Graduate and Doctoral Students Use in Reading Both Chinese and English Academic Publications*. Paper presented at the 10th conference on English Teaching and Learning in the Republic of China. (Published in Mandarin Chinese).



大學生英文閱讀年級指數的探討-單字標準化測驗的應用研究

鄭宗諭

中國醫藥大學 通識教育中心 副教授

摘 要

本研究利用單字標準化測驗探討大學生的英文閱讀年級指數。目的是要找出國內大學生的英文閱讀能力大約同等於美國學制的那一年級。145 位來自一所醫學大學的學生在期中考時接受 The Slossen Oral Reading Test 的測試。敘述性統計學分析結果：全部受測者的年級分佈為小學一年級到六年級。詳細分佈比率如下：六年級(4.8%)、五年級(26.2%)、四年級(38.6%)、三年級(24.15%)、二年級(5.5%)、一年級(0.7%)。眾數及平均年級為四年級。本研究的對象為非醫學類學生，研究結果只能適用於中、西、牙醫、藥學及物理治療以外的其他科系學生。如何利用本研究結果協助分班以及指定配合學生能力的英語教材，在文中有詳細的討論。

關鍵詞：英文閱讀測驗、標準化測驗、能力分級、閱讀年級指數

