English Language Teaching; Vol. 11, No. 11; 2018
ISSN 1916-4742 E-ISSN 1916-4750
Published by Canadian Center of Science and Education

Item Types: Their Effect on the Sensitivity of Multiple-Choice Cloze Tests

Abby Deng-Huei Lee[1]

[1] Department of Applied Foreign Languages, College of Humanities & Applied Information, Chia Nan University of Pharmacy & Science, Tainan, Taiwan

Correspondence: Abby Deng-Huei Lee, Chia Nan University of Pharmacy and Science, No. 60, Section 1, Erren Road, Rende District, Tainan City 71710, Taiwan. E-mail: abbydhlee@mail.cnu.edu.tw

Received: September 5, 2018 Accepted: October 11, 2018 Online Published: October 13, 2018

doi: 10.5539/elt.v11n11p98 URL: http://doi.org/10.5539/elt.v11n11p98

Abstract

To evaluate the sensitivity of multiple-choice cloze (MCC) tests that use different types of items—syntactic, semantic, and connective—to assess reading ability, 170 English as a foreign language (EFL) students in a vocational college in Taiwan were recruited. The students were divided into two groups (level A and level B) based on their scores on 4 classroom reading comprehension tests. Both groups then took 9 MCC tests that included a total of 50 cloze questions. Connective items were most sensitive for assessing reading ability. Research results and pedagogical applications are discussed.

Keywords: cloze item type, connective clue, syntactic clue, semantic clue, multiple-choice cloze (MCC) test, English as a foreign language (EFL)

1. Introduction

The role of cloze tests as a measuring tool of overall reading ability has been controversial. One main reason for this, among others, is that when counting the scores of cloze test, we often ignore the discrepant traits of cloze items. Thus I study the construct validity of multiple-choice cloze (MCC) test scores. Specifically, I analyzed syntactic, semantic, and connective clues by comparing the MCC test scores of two different levels of students.

1.1 Statement of the Problem

The word "cloze" is derived from "closure", a term that Gestalt psychologists use to describe the human self-organizing tendency to form a whole. Cloze has been a hot issue in language testing and teaching since it was developed (Taylor, 1953). Interestingly enough, cloze is still unfamiliar to people outside the EFL field, and even some older editions of *Microsoft Word*'s spell checker do not recognize the word "cloze" and want to replace it with "close". Over the past six decades, the most commonly discussed problem with cloze is its sensitivity: Can cloze test long-range constraints and so be able to measure integrative language proficiency or reading comprehension, or does it just test a single component of language at a time?

Some important pro-cloze studies are (Taylor, 1953; Cziko, 1978; Oller, 1979; Bachman, 1982, 1985; Jonz, 1990, 1991; Hale, Stansfield, Rock, Hicks, Butler, & Oller, 1988; Abraham & Chapelle, 1992; Wu, 1994; Chatel, 2001; Kobayashi, 2002; Ravand & Sardai, 2017). In contrast, some important skeptical or anti-cloze studies are (Alderson, 1979; Brown, Yamashiro, & Ogane, 1999; Shanahan & Kamil, 1982; and Ashby-Davis, 1985).

1.1.1 Main Reason for the Disagreement

The main reason for the disagreements about cloze studies is inconsistent results because the discrepant traits of the types of cloze items have been ignored. Traditionally, to derive the score of a cloze passage is to calculate the unweighted sum of the scores of all its items. Bachman (1985) emphasized that "although research on the cloze test has offered differing evidence regarding what language abilities it measures, there is a general consensus among researchers that not all the deletions in a given cloze passages measure exactly the same abilities" (p. 535). If it is true that the items do not always measure the same dimension, then the total score would not always yield consistent results. This is why readers must be more cautious when explaining the scores of a cloze passage: the discrepancies of the sensitivity of cloze tests might occur because of the deletion of different types of items. To explain the outcomes of cloze tests, readers must consider the traits of each cloze item because different types of cloze items affect the passage in different ways.

1.1.2 Classification of Cloze Items

There are several ways to classify cloze items. Bensoussan and Ramraz (1982, p. 12) assigned them to two categories: the micro- and the macro-level. Lee (1985) concluded that constraints on cloze might derive from just two sources: relative "closedness" as opposed to relative "openness". Bachman (1982) divided cloze items into three: 1) syntactic clues, which depend upon only clause-level context, 2) cohesive clues, which depend upon the interclausal or intersentential cohesive context, and 3) strategic (semantic) clues, which depend upon parallel patterns of coherence. Three years later, Bachman (1985) revised his original categories into four: "1) within clause; 2) across clauses, within sentence; 3) across sentences, within text; and 4) extratextual" (p. 539). Jonz (1990) divided the first of Bachman's 1985 categories into two ([1] clausal-level syntax [Category 1-S] and [2] clausal-level lexis [Category 1-L]), and created a 5-category classification. Many cloze studies discuss the differences between content and function words (e.g., Abraham & Chapelle, 1992; Kobayashi, 2002). Hale et al. (1988) subdivided the MCC items into four different categories based on content analysis: 1) Reading Comprehension/Grammar (RG), 2) Reading Comprehension/Vocabulary (RV), 3) Grammar/Reading Comprehension (GR), and 4) Vocabulary/Reading Comprehension (VF) (Table 1).

The classification of types of cloze items stated above is used for conventional blank-filling cloze tests, but the present study focuses on MCC tests. A three-part classification was used in the present study: syntactic, semantic, and connective clues. Briefly, syntactic clues refer to within-clause grammatical points, semantic clues refer to lexical points, and connective clues refer to both across-clauses and across-sentences conjunctions and to transitional words and phrases (Table 2).

Table 1. Classification of types of cloze items

Author/Researcher	**Classification of Cloze Types**	**Description**
Bachman, 1982	Three-part system:	
	1. Syntactic	Depends only upon clausal-level context.
	2. Cohesive	Depends upon the interclausal or intersentential cohesive context.
	3. Strategic	Depends upon parallel patterns of coherence.
Bachman, 1985	Four-part system:	
	1. Within Clause	Bachman revised his 1982 three-part system into a four-part system based on the four levels of context
	2. Across Clauses, Within Sentence	(Bachman, 1985, p. 539).
	3. Across Sentences, Within Text	
	4. Extratextual	
Bensoussan, 1983	Three-part system:	
	1. Form	Form: verb, noun, pronoun, determiner, adjective, etc.
	2. Content	
	3. Form + Content	
Bensoussan & Ramraz, 1982	Two-part system:	
	1. Micro	Micro constraints: 1-S, 1-L, & category 2 in Jonz, 1990, p. 73.
	2. Macro	Macro constraints: categories 3 & 4 in Jonz, 1990, p. 73.
Hale et al., 1988:	Four-part system:	

Author/Researcher	Classification of Cloze Types	Description
	1. RG	Reading Comprehension/Grammar: Taps propositional information at an interclausal level and emphasizes knowledge of syntax.
	2. RV	Reading Comprehension/Vocabulary: Requires long-range constraints and a lexical choice.
	3. GR	Grammar/Reading Comprehension: Taps knowledge of surface syntax and within-clause proposional information.
	4. VF	Vocabulary/Reading Comprehension: Primarily deals with vocabularyand invokes reading comprehension within clause boundaries.
Jonz, 1990:	Five-part system:	
	1.Clause Level Syntax (1S)	Jonz divided Bachman's (1985) 1st category (Within-Clause) into 2 (1S & 1L) and thus created a five-part system.
	2. Clause Level Lexis (1L)	
	3. Across Clauses, Within Sentence	
	4. Across Sentences, Within Text	
	5. Extratextual	
Kobayashi, 2002;	Two-part system:	
Abraham & Chapelle, 1992	Function Words	Prepositions, articles, conjunctions, pronouns, determiners, etc.
	Content Words	Verbs, nouns, adjectives, adverbs, etc.
Lee 1985: p. 139	Two-part system:	
	1.Relative Closedness	Based on short-range constraints.
	2.Relative Openness	Based on long-range constraints.

Table 2. Similarity between the categories of the present study and some main studies in Table 1 above

Present study	Bachman, 1982	Bachman, 1985	Bensoussan & Ramraz, 1982	Hale et. al., 1988	Jonz,1990	Kobayashi, 2002; Abraham & Chapelle, 1992
Syntactic	**Syntactic	**Within Clause	**Micro	*RG	**Local Syntax (1S),	**Function Words
	Cohesive	Across Clauses, Within Sentence	Macro	RV	Local Lexis (1L),	Content Words
	Strategic	Across Sentences, Within Text		**GR	Across Clauses, Within Sentence	

		Extratextual		VF	Across Sentences, Within text	
					Extratextual	
Semantic	Syntactic	**Within Clause	*Micro	RG	Local Syntax (1S),	Function Words
	Cohesive	Across Clauses, Within Sentence	Macro	*RV	**Local Lexis (1L),	**Content Words
	**Strategic	Across Sentences, Within Text		GR	Across Clause, Within Sentence	
		Extratextual		**F	Across Sentences, Within text	
					Extratextual	
Connective	Syntactic	Within Clause	Micro	*RG	Local Syntax (1S),	**FunctionWords
	**Cohesive	**Across Clauses, Within Sentence	**Macro	RV	Local lexis (1L),	*Content Words
	*Strategic	**Across Sentences, Within text		**GR	**Across Clauses, Within Sentence	
		Extratextual		VF	**Across Sentences, Within text	
					Extratextual	

Note. **Identical or very similar; *only slightly similar.

1.2 Review of Related Literature

Some studies have investigated the concurrent validity of cloze and reported data to explain the sensitivity of cloze items. These studies show that cloze item types are important for explaining the cloze scores.

Two important questions related to the focus of the present study will be briefly reviewed: (1) the different dimensions and functions of different types of cloze items, and (2) the association between cloze performance and student proficiency level.

Several studies have emphasized that different types of cloze items address different dimensions of language points. Bachman (1985) said that "not all words in a given text function at only one or at the same structural level, and it therefore seems unreasonable to expect all deletions to depend equally on the same level or range of context for closure" (p. 538). Cziko (1978) claimed that "the contextual information available to the reader can be of three types: syntactic constraints, semantic constraints and discourse constraints" (p. 473). Syntactic constraints are those provided by the preceding words and the syntactic rules of the language (e.g., in the sentence *The boy ... in the snow.* the word *The* will be followed by a noun). Semantic constraints are those provided by the meaning of the preceding words (e.g. the words *The boy* at the beginning of a sentence will most likely be followed by a verb phrase describing something a boy is likely to do), and discourse constraints are those provided by the topic of the text (e.g., all the sentences in a reading about skiing will be in some way related to skiing) (p. 473). Jonz (1991) stated that "…cloze was sensitive to features of textual processing that ranged far beyond the narrow focus of local syntactic constraint" (p. 6). Chihara, Oller, Weaver, and Chavez-Oller (1994) pointed out that "short-range constraints are often confused with mere syntactic elements, and yet they often involve lexical connections that transcend traditional conceptions of phrase structure syntax"

(p. 144). Abraham and Chapelle (1992) reported that "in the multiple-choice cloze the potential answers were limited…Thus the multiple-choice format made content words differentially easier than function words" (p. 472).

In addition, some studies emphasize the close relationship between student proficiency levels and cloze test scores. Bachman (1985) found that "the percentage of correct closures was higher for groups with higher language proficiency than for groups with lower proficiency" (p. 536). Chihara et al. (1994) found that "as subjects increase in proficiency, they become more able to benefit from discourse constraints (i.e. long-range constraints) ranging across sentence boundaries" (p. 142).

The present study aimed to answer this question: Are there language proficiency-based differences in student scores on MCC tests that use different types of cloze items (syntactic, semantic, and connective)?

2. Methods

2.1 Participant Characteristics

About 200 students were recruited. Most had completed 6 years of EFL classes in junior and senior high school, and about half of them had also finished one year of 2-hour/week English classes in their first year of college. The students took four 15-minute 20-question reading comprehension quasi-placement tests during the first four weeks of the study. Students were then grouped into two levels: Level A students were those who had correctly answered at least 14 of the 20 questions in the 4 reading comprehension tests, while Level B students were those who had correctly answered 7-13 questions. Students who had correctly answered fewer than 7 questions were excluded from analysis because their tests had little reference value. Finally, 170 students completed the study.

2.2 Materials

The 4 reading comprehension passages used in this study to determine student language proficiency levels were taken from Taiwan's Testing Center for Technological and Vocational Education (TCTE) website (http://www.tcte.edu.tw/down_exam.php). Each passage was 180-240 words, and the difficulty levels of the passages were between Flesch-Kincaid Grade Levels 8.4 and 11.6, according to the *Automatic Readability Checker* (http://www.readabilityformulas.com/free-readability-formula-tests.php).

The 9 cloze passages were also from the TCTE. Each passage was 160-201 words; 5-7 items followed each passage, and the difficulty levels of the passages were between Flesch-Kincaid Grade Levels 8.3 and 11.5.

2.3 Procedures

Students who had taken the required one year of 2-hour/week English classes in their first year of college were randomly chosen for the study. At the outset of the experiment, all students took a reading comprehension test randomly chosen from the item bank of the entrance examination of Vocational and Technical Education in Taiwan. Passages were of comparable length and difficulty level (about 180-240 words); difficulty levels were between Flesch-Kincaid Grade Levels 9.5 and 11.6, according to the *Automatic Readability Checker*.

Then in the following four weeks, each student answered 10-15 MMC test questions per week in 15 minutes (a total of 50 cloze items). The 50 MCC test questions were graded and categorized into three types of cloze item clues by 5 experienced EFL instructors blinded to the identities of the tested students. If the agreement of the 5 EFL instructors was less than 80%, the disagreements were discussed and reset. Three dubious and complicated items were discarded and replaced with new answers. Consequently, 20 of the 50 items were syntactic, 14 were semantic, and 16 were connective. Unequal variance *t* tests were used to investigate the differences between the three different item types. The independent variables of the study were Level A and Level B, and the dependent variables were the scores of the three different item types.

3. Results

The unequal variance *t* test was used to determine the differences between the test scores of Levels A and B. In Table 3, for syntactic item types, the *t* test value was 2.740 with 115 degrees of freedom; for semantic item types, the *t* test value was 4.902 with 113 degrees of freedom; and for connective item types, the *t* test value was 7.493 with 130 degrees of freedom. All three differences were significant ($p < 0.05$). These results seem to be natural and reasonable, since there is difference of the language proficiency between Level A and Level B students. The mean differences of the three different item types (Level A minus Level B) were then compared.

In Table 4, the mean difference between syntactic and semantic item types was not significant (*t* test value was 0.809); however, the mean differences between connective and syntactic item types and between connective and semantic item types were significant (*t* test values were 2.796 and 2.315, respectively). Results show that connective clues were the most sensitive.

Table 3. Descriptive statistics

Variables	Level A	Level B	H_0: $\mu_A=\mu_B$
	(mean ± SD)	(mean ± SD)	*t* test
Syntactic	9.355 ± 2.940	8.072 ± 2.843	2.740**
Semantic	6.779 ± 2.274	5.018 ± 2.144	4.902**
Connective	9.847 ± 2.398	6.837 ± 2.661	7.493**
n (sample size)	59	111	

Table 4. *t* test for mean difference

	Level A MINUS Level B		
Variables	Mean Difference ($\Delta\mu_{A-B}$) ± SD		H_0: $\Delta\mu_{1,A-B}= \Delta\mu_{2,A-B}$
			t test
Syntactic	1.283 ± 0.468	Syntactic vs. Semantic	0.809
Semantic	1.761 ± 0.359	Connective vs. Syntactic	2.796**
Connective	3.010 ± 0.401	Connective vs. Semantic	2.315**

4. Discussion

4.1 Findings

The study's most important findings were that the connective clues (i.e., across-clause, within sentence clues and across-sentence, within-text clues) were the most sensitive indicators of language level and reading comprehension: the differences between connective clues and the other two types (syntactic and semantic) were statistically significant. There was, however, no statistically significant difference between syntactic and semantic clues. Although more work remains to be done, these findings suggest that as language proficiency increases, students better understand the discourse constraints across sentence boundaries. These findings confirm Bachman's (1985) statement that "to develop a test that could potentially measure textual relationships beyond the clausal level, it was necessary to identify criteria for classifying and selecting words to be deleted" (p. 538). These results are also consistent with some other studies (e.g., Chihara et al., 1994; Jonz, 1990, 1991). They also support the pro-cloze hypothesis that cloze can measure integrated reading comprehension in addition to testing discrete points of language use. The findings of the study might help clarify the nature of cloze items and enable researchers, test makers, and test administrators to rationally rather than randomly choose which words and phrases should be deleted. This kind of knowledge should lead to better language tests and to assessments that are more accurate.

4.2 Limitations

This study has some limitations. First, because of the homogeneity of the participants, readers must be cautious when interpreting the outcomes and conclusions. Second, only a three-part system was used, not the four- or five-part systems that some others (e.g., Bachman, 1982, 1985; Jonz, 1990, 1991) used. Third, the five experienced EFL teachers in this study reached only 80% agreement when judging the three item types. These item types might need to be redefined in future studies to make the findings more general and convincing. Fourth, because of the student homogeneity and the narrow range of difference in language proficiency levels in this study, the two levels are somewhat arbitrary. In future studies, a broader range of difference in language proficiency levels would likely be more accurate.

4.3 Pedagogical Application of the Study

Several useful conclusions related to EFL teaching and testing can be drawn from the present study's findings. First, the study supports the pro-cloze position that cloze, including the MCC test, can be used to measure global and integrated reading proficiency. Second, to increase reading ability proficiency, it is important to develop the reader's ability to grasp long-range contextual constraints. Third, when designing a rational-deletion cloze test, connective clues cannot be neglected. Fourth, when designing or interpreting MCC tests, EFL teachers should carefully choose item types. Fifth, experienced EFL teachers need to be involved when designing a rational deletion MCC test. Sixth, properly constructed MCC tests can teach some foreign language students who are

characterized as word-by-word readers to perceive how individual words relate to one another and, therefore, to pay attention to contextual constraints to promote their integrated reading proficiency.

4.4 Recommendations for Future Research

The language levels of the participants in this study were low-to-intermediate and tended to be homogeneous. If advanced-level students are recruited for future studies, the outcomes and conclusions will be more convincing.

In addition, in similar studies, if four- or five-part systems (e.g., Bachman, 1985; Jonz, 1990; Hale et al., 1992) are used and can be factor analyzed to ascertain whether cloze procedure scores can account for a model that reflects the nature of the cloze item types (Jonz, 1990), the outcomes and conclusions should be more meaningful. Because MCC tests are currently commonly used in national entrance examinations in many countries and in internationally famous English proficiency examinations like TOEFL and TOEIC (Lee & Wu, 2018, p. 2), the research on MCC item traits should be meaningful.

Acknowledgments

I am grateful to Professor S. T. Yu (National Tsing Hua University, Hsinchu, Taiwan) for guiding the statistical analyses, and to Professor Richard Jenn-Rong Wu (Chia Nan University of Pharmacy) for offering his experience in cloze study.

References

Abraham, R. G., & Chapelle, C. A. (1992). The meaning of cloze test scores: An item difficulty perspective. *The Modern Language Journal, 76*(4), 468-479. https://doi.org/10.1111/j.1540-4781.1992.tb05394.x

Ajiden P., Yaghoubi-Notash, M., & Khalili, A. (2017). An Investigation of the Learning Strategies as Bias Factors in Second Languge Cloze Tests. *Advances in Language and Literary Studies, 8*(2), 91-100. https://doi.org/10.7575/aiac.alls.v.8n.2p.91

Alderson, J. C. (1979). The cloze procedure and proficiency in English as a foreign language. *TESOL Quarterly, 14*(2), 219-223. https://doi.org/10.2307/3586211

Ashby-Davis, C. (1985). Cloze and comprehension: A qualitative analysis and critique. *Journal of Reading, 28*(7), 585-589.

Automatic Readability Checker. (n. d.). Retrieved from http://www.readabilityformulas.com/free-readability-formula-tests.php

Bachman, L. F. (1982). The trait structure of cloze test scores. *TESOL Quarterly, 16*(1), 61-70. https://doi.org/10.2307/3586563

Bachman, L. (1985). Performance on cloze tests with fixed-ration and rational deletions. *TESOL Quarterly, 19*(3), 535-556. https://doi.org/10.2307/3586277

Baghaei, P. & Ravand, H. (2016). Modeling Local Item Dependence in Cloze and Reading Comprehension Test Items Using Testlet Response. *Psicologica, 37*, 85-104.

Barnett, M. (1986). Syntactic and lexical/semantic skill in foreign language reading: Importance and interaction. *The Modern Language Journal, 70*(4), 343-349. https://doi.org/10.1111/j.1540-4781.1986.tb05286.x

Bensoussan, M., & Ramraz, R. (1982). The fill-in test: Combining modified cloze techniques to test English as a foreign language. *ERIC Reports, ED 227 652; FL 013 4784*, 1-68.

Bensoussan, M. (1983). Multiple-choice modifications of the cloze procedure using word-length and sentence-length blanks. *Foreign Language Annals, 16*(3), 189-198. https://doi.org/10.1111/j.1944-9720.1983.tb01448.x

Bensoussan, M., & Ramraz, R. (1984). Testing EFL reading comprehension using a multiple-choice rational cloze. *The Modern Language Journal, 68*(3), 230-239. https://doi.org/10.1111/j.1540-4781.1984.tb01569.x

Brown, J. D., Yamashiro, A. D. & Ogane, E. (1999). Tailoring cloze: Three ways to improve cloze tests. *University of Hawaii Working Papers in ESL, 17*(2), 107-131.

Cameron, C. A. (1987). *Automated cloze procedures as research and Teaching Tool*. Paper presented at the ANNUAL Meeting of the Canadian Psychological Association, Vancouver, British Columbia, Canada. June 1987.

Chatel, R. G. (2001). Diagnostic and instructional uses of the cloze procedure. *New England Reading Association Journal, 37*(1), 3-8.

Chihara, T., Oller, J. W., Weaver, K., & Chavez-Oller, M. A. (1994). Are cloze items sensitive to constraints across sentences? In J. W. Oller, & J. Jonz (Eds.) *Cloze and coherence* (pp. 135-148). Cranbury, New Jersey: Bucknell University Press.

Cziko, G. A. (1978). Differences in first- and second-language reading: The use of syntactic, semantic and discourse constraints. *The Canadian Modern Language Review, 34,* 473-489. https://doi.org/10.3138/cmlr.34.3.473

Greene, B. (2001). Testing reading comprehension of theoretical discourse with cloze. *Journal of Research in Reading, 24*(1), 82-98. https://doi.org/10.1111/1467-9817.00134

Hale, G. A., Stansfield C. W., Rock, D. A., Hicks, M. M., Butler, F. A., & Oller, J. W. Jr. (1988). Multiple-choice cloze items and the test of English as a Foreign Language. *TOEFL Research Report 26*. Educational Testing Service, Princeton, NJ.

Jonz, J. (1990). Another turn in the conversation: What does cloze measure? *TESOL Quarterly, 24*(1), 61-83. https://doi.org/10.2307/3586852

Jonz, J. (1991). Cloze item types and second language comprehension. *Language Testing, 8*(1), 1-22. https://doi.org/10.1177/026553229100800102

Kobayashi, M. (2002). Cloze tests revisited: exploring item characteristics with special attention to scoring methods. *The Modern Language Journal, 86*(4), 571-586. https://doi.org/10.1111/1540-4781.00162

Lee, A. D. H., & Wu, R. J. R. (2018). An instructional application of the multiple-choice cloze: A case study in the EFL classroom. *English Language and Literature Studies, 8*(2), 1-9. https://doi.org/10.5539/ells.v8n2p1

Lee, Y. P. (1985). Investigating the validity of the score. In Y. P. Lee, A. Fok, R. Lord, & G. Low (Eds.), *New directions in language testing* (pp. 137-147). Oxford: Pergamon Press.

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

Ravand, H., & Sardari, S. A. (2017). The effect of gender and academic background on cloze and reading comprehension performance using G-theory. *Psychological Test and Assessment Modeling, 59*(2), 207-218.

Slabakova, R. (2006). Is there a critical period for semantics? *Second Language Research, 22*(3), 302-338. https://doi.org/10.1191/0267658306sr270oa

Shanahan, T. Kamil, M. L. & Tobin, A. W. (1982). Cloze as a measure of intersentential comprehension. *Reading Research Quarterly, 17*(2), 229-255. https://doi.org/10.2307/747485

Spolsky, B. (2000). Language testing in The Modern Language Journal. *The Modern Language Journal, 84*(4), 536-552. https://doi.org/10.1111/0026-7902.00086

Taylor, W. L. (1953). Cloze procedure: A new tool for measuring readability. *Journalism Quarterly, 30*, 415-433. https://doi.org/10.1177/107769905303000401

The Testing Center for Technological and Vocational Education in Taiwan. (n. d.). Retrieved from http://www.tcte.edu.tw/down_exam.php

Williams, R. S., Ari, O., & Santamaria, C. N. (2011). Measuring College Students' Reading Comprehension Ability Using Cloze Tests. *Journal of Research in Reading, 34*(2), 215-231. https://doi.org/10.1111/j.1467-9817.2009.01422.x

Wu, R. J. R. (1994). *The sensitivity of the cloze procedure to discourse constraints in English as a foreign language.* Unpublished Doctoral Dissertation, the University of Kansas, Lawrence, KS, USA.

Xiao, X. Y., & Ho, C. (2014). Weakness in semantic, syntactic and oral Language expression contribute to reading difficulties in Chinese dyslexic children. *View Issue TOC, 20*(1), 74-98. https://doi.org/10.10021/dys.1460

Appendix A

1. An example of the reading comprehension tests:

(Source: Item Banks of The Testing Center for Technological and Vocational Education in Taiwan. [n. d.])

Retrieved from (http://www.tcte.edu.tw/down_exam.php)

An example of the Reading Comprehension Passage:

Wasting food is morally wrong across cultures. After all, about 800 million people worldwide suffer from hunger. But according to a report from the United Nations, people around the world waste 2.9 trillion pounds of food a year. In developing nations, much food is lost after the harvest for lack of storage facilities, good roads, and refrigeration. In comparison, developed nations waste more food when retailers order, serve, or display too much and when consumers ignore leftovers in the fridge or get rid of foods before they have expired. Wasting food also negatively affects our environment. If we produce food that no one eats, we also waste the water, fertilizer, seeds, fuel, and land needed to grow it. For example, a year's global production of uneaten food uses as much water as the annual flow of the Volga River, Europe's largest river. In the U.S., retailers and consumers are estimated to discard 133 billion pounds of food each year. These examples just show how many natural resources have been wasted, and how our environment might be damaged by air or river pollution. Reducing food waste has become an important issue that needs to be dealt with. Some U.S. schools are setting up sharing tables, letting students serve themselves portions they know they will eat, and giving them more time for lunch. Many businesses, such as grocery stores and restaurants, have planned to order food wisely and give excess food to charities. It is also important to examine conditions farther up the supply chain, where supermarket standards and ordering methods lead to massive food waste.

(*265 words, Difficulty level: Flesch-Kincaid Grade Level: 8.4*)

1. Which of the following is NOT the reason why the developing nations lose much food?

(A) Their roads are in bad condition. (B) The refrigerators are not enough.

(C) They have a problem of air pollution. (D) They lack facilities to store their food.

2. What is the main idea of the second paragraph?

(A) Some moral issues of food waste. (B) The suggestions for preserving food.

(C) The statistics of consumers' food waste. (D) Some environmental impacts of food waste.

3. Which of the following is closest in meaning to discard in the second paragraph?

(A) turn in (B) look over (C) throw away (D) make up

4. According to this passage, which of the following is the method of solving the problem of food waste?

(A) Offering excess food to charities. (B) Asking students to bring food to school.

(C) Urging supermarkets to give food to poor families. (D) Encouraging retailors to buy a large amount of food.

5. Which of the following is true according to the passage?

(A) Developing nations waste more food than developed nations. (B) The supply chain is one of the major problems of food waste.

(C) Consumers usually buy as much food as their families would need. (D) About 133 billion pounds of food are wasted each year in the world.

2. Examples of cloze item types (in *italics*):

[1]

In my free time, I really like surfing the Internet. When I get home from work, I turn on my computer, wait until it boots up completely, and then go online. With a broadband Internet service, webpages load faster with a high-speed __1__ , and I can upload and download files faster. I usually check my e-mail first and write a few messages to family and friends. I sometimes __2__ the local news headlines at my favorite news website and read up on the __3__ local and international news. This website often provides video news clips that I can view online. I sometimes order products online to save money and time __4__ going to a store and buying what I am looking for. Whatever I do, I realize __5__ there are problems with using the Internet, including scams, identity theft, and viruses. I am very careful not to give out my personal information. Using the Internet can be a fun and convenient way of shopping and finding out new information, but I just need to be careful.

(*177 words; Difficulty Level: Flesch-Kincaid Grade Level: 8.3*)

1. (A) description (B) *confusion* (C) *connection* (D) dedication

Item type: *semantic clue*

2. (A) put (B) *scan* (C) direct (D) mix

Item type: *semantic clue*

3. (A) *latest* (B) latter (C) less (D) lest

Item type: *syntactic clue*

4. (A) *instead of* (B) inclusive of (C) in sense of (D) in spite of

Item type: *connective clue*

5. (A) what (B) whatever (C) which (D) *that*

Item type: *syntactic clue*

[2]

Most people have heard the old English saying, "An apple a day keeps the doctor away." This means that eating an apple every day can keep people ___6___ . But is it really true that eating apples will help people keep in good physical shape? ___7___ research studies, the answer is yes. The results of these studies suggest that apples may play an important role in reducing the risk of a wide variety of diseases ___8___ heart problems and lung cancer. In addition, apples are sweet, low in calories and totally fat-free. Many health problems are associated ___9___ people being overweight. Eating an apple can help you satisfy your hunger with lower calories, and consequently stay fit. There are many other advantages that apples can provide to your health including improving memory, boosting the immune system, and so on. ___10___ , apples are not only delicious to eat, but also beneficial to your body. So, why not begin your an-apple-a-day habit today?

(*162 words; Difficulty Level: Flesch-Kincaid Grade Level: 8.4*)

6. (A) heal (B) heals (C) *healthy* (D) healthily

Item type: *semantic clue*

7. (A) *According to* (B) In fact (C) By contrast (D) So that

Item type: *connective clue*

8. (A) instead of (B) *such as* (C) less than (D) after all

Item type: *connective clue*

9. (A) for (B) from (C) about (D) *with*

Item type: *syntactic clue*

10. (A) In order to (B) *In short* (C) In case (D) In the beginning

Item type: *connective clue*

Appendix B

Table. 5 The Raw Scores of the Cloze Items and Students' Reading Comprehension Levels

L #	L	Syn. n = 20	Sem. n = 14	Con. n = 16	L #	L	Syn. n = 20	Sem. n = 14	Con. n = 16	L #	L	Syn. n = 20	Sem. n = 14	Con. n = 16
A1	A	5	6	11	B1	B	8	6	7	B60	B	7	6	8
A2	A	7	8	11	B2	B	10	4	9	B61	B	3	1	8
A3	A	12	10	13	B3	B	10	6	5	B62	B	6	1	6
A4	A	13	6	12	B4	B	5	8	6	B63	B	8	6	5
A5	A	7	5	6	B5	B	5	4	6	B64	B	7	7	6
A6	A	12	7	9	B6	B	10	2	9	B65	B	11	4	6
A7	A	9	6	7	B7	B	9	4	3	B66	B	7	6	6
A8	A	6	6	7	B8	B	10	6	6	B67	B	8	5	5
A9	A	11	5	11	B9	B	7	6	10	B68	B	10	6	10
A10	A	14	9	9	B10	B	11	6	10	B69	B	5	3	7
A11	A	9	5	8	B11	B	8	4	9	B70	B	6	5	5

L #	L	Syn. n = 20	Sem. n = 14	Con. n = 16	L #	L	Syn. n = 20	Sem. n = 14	Con. n = 16	L #	L	Syn. n = 20	Sem. n = 14	Con. n = 16
A12	A	6	5	6	B12	B	10	7	6	B71	B	7	4	6
A13	A	7	2	11	B13	B	10	4	7	B72	B	8	8	7
A14	A	10	8	11	B14	B	7	4	8	B73	B	6	3	7
A15	A	8	6	9	B15	B	6	2	6	B74	B	11	5	4
A16	A	11	11	13	B16	B	6	5	7	B75	B	4	4	7
A17	A	9	8	10	B17	B	7	5	7	B76	B	7	7	6
A18	A	12	8	13	B18	B	11	3	5	B77	B	10	5	4
A19	A	7	6	6	B19	B	7	6	5	B78	B	11	4	8
A20	A	9	4	8	B20	B	10	6	6	B79	B	6	5	8
A21	A	8	6	11	B21	B	5	3	4	B80	B	9	6	7
A22	A	10	7	11	B22	B	12	8	7	B81	B	18	13	8
A23	A	14	6	6	B23	B	11	6	7	B82	B	6	6	6
A24	A	11	9	12	B24	B	6	5	4	B83	B	6	6	8
A25	A	7	7	10	B25	B	5	5	3	B84	B	6	3	5
A26	A	10	7	8	B26	B	7	3	7	B85	B	11	6	8
A27	A	16	7	11	B27	B	6	4	4	B86	B	12	6	9
A28	A	7	5	6	B28	B	5	2	4	B87	B	6	2	7
A29	A	8	4	9	B29	B	10	2	5	B88	B	14	4	11
A30	A	6	4	7	B30	B	5	6	2	B89	B	7	6	8
A31	A	9	6	12	B31	B	9	3	6	B90	B	9	6	10
A32	A	11	6	11	B32	B	5	1	6	B91	B	4	7	7
A33	A	11	6	8	B33	B	6	7	6	B92	B	7	2	9
A34	A	6	6	11	B34	B	10	5	9	B93	B	9	6	11
A35	A	11	12	1`2	B35	B	5	2	5	B94	B	7	5	10
A36	A	8	8	9	B36	B	10	7	2	B95	B	8	5	9
A37	A	5	5	8	B37	B	8	5	7	B96	B	7	6	10
A38	A	8	7	10	B38	B	3	2	3	B97	B	10	5	9
A39	A	12	12	14	B39	B	12	7	7	B98	B	12	5	9
A40	A	14	12	13	B40	B	10	3	7	B99	B	9	6	12
A41	A	8	5	10	B41	B	8	2	4	B100	B	9	4	6
A42	A	11	5	11	B42	B	8	3	5	B101	B	8	2	5
A43	A	10	7	12	B43	B	6	7	9	B102	B	8	5	5
A44	A	9	5	8	B44	B	7	3	4	B103	B	9	7	6
A45	A	7	6	11	B45	B	8	9	5	B104	B	5	5	7
A46	A	9	6	9	B46	B	11	4	7	B105	B	8	7	6
A47	A	7	8	10	B47	B	7	5	9	B106	B	10	7	6
A48	A	8	6	9	B48	B	4	3	6	B107	B	8	8	9
A49	A	14	10	12	B49	B	9	7	7	B108	B	16	11	16
A50	A	5	2	11	B50	B	3	5	6	B109	B	17	9	15
A51	A	11	4	9	B51	B	6	3	4	B110	B	12	7	15

L #	L	Syn. n = 20	Sem. n = 14	Con. n = 16	L #	L	Syn. n = 20	Sem. n = 14	Con. n = 16	L #	L	Syn. n = 20	Sem. n = 14	Con. n = 16
A52	A	11	9	12	B52	B	5	7	5	B111	B	15	11	16
A53	A	9	9	7	B53	B	7	1	8					
A54	A	14	8	11	B54	B	4	4	5					
A55	A	9	5	6	B55	B	5	4	4					
A56	A	7	5	6	B56	B	9	5	1					
A57	A	5	6	7	B57	B	7	5	5					
A58		15	11	16	B58	B	10	3	5					
A59		17	10	14	B59	B	5	4	4					

Note. L: level; Level A: 14 -20; Level B: 7-13; syn.: syntactic clues; sem.: semantic clues; con.: connective clues.