

A Comparative Study on C-test vs. Cloze Test as Tests of Reading Comprehension

Parviz Ajideh, Sorayya Mozaffarzadeh*

Department of English Language and Literature, University of Tabriz, Iran

ABSTRACT

Since its introduction to the testing world by Taylor in 1953, cloze tests have been widely used for measuring reading comprehension. But in 1982, Klein and Braley by criticizing cloze procedure, they introduced their newly developed testing procedure, C-test, which resolved the cloze tests deficiencies. After that, C-test has been used for measuring reading comprehension and language proficiency while cloze procedure also has continued his popularity. However, little research has done on the effectiveness of these two tests as reading tests. The present study aims to investigate the effectiveness of cloze test vs. C-test as measures of reading comprehension. To this end, one traditional C-test and one fixed ratio (n=7) multiple-choice cloze test prepared from reading passages with similar readability level. The subjects of the study were 27 female EFL advanced learners. Contrary to the finding of previous research, the results of the study revealed that multiple-choice cloze is a better measure of reading comprehension. The estimated correlation between these tests was 0.38. The implications of the finding and suggestions for more studies are discussed within a foreign language testing context.

KEYWORDS: C-test, cloze test, reading comprehension

1. INTRODUCTION

1.1 Testing reading comprehension

Reading is probably the most common of the four skills to be tested, and it may seem to be the easiest of the skills to test. However, testing reading does have difficulties, and there are issues that anyone testing reading should be aware of. How the ability to read text in a foreign language might be assessed in a best way has long interested language testing researchers. In English as a foreign/second language reading comprehension tests, it is very common to include a series of related items that are based on the same reading passage [1]. These items can be placed after a passage, as in traditional comprehension questions, multiple-choice or short-answer or embedded in the passage itself as in cloze or c-test [2]. Of important issues discussed in reading comprehension tests one can name text selecting and methods used for testing reading.

One of the more complex tasks facing in the testing of reading comprehension especially in English as a foreign language (FEL) is the selection of appropriate reading text. Day [3] discusses seven factors which should be considered in the selection of texts. The most important factors concerning testing reading include: readability, cultural suitability and appearance.

The factor of readability of the text is one of the significant considerations in the selecting of the reading texts especially for EFL students (for more details refer to Day 1994). If a reader-text mismatch, the result will be the user failing to use or ignoring the text [4]. To avoid such mismatch, educators would like a tool to check if a given text would be readable by its intended audience. Inventing such tools has been the primary focus of readability research for the past 90 years (Kondru, 2006 cited in [4]). To this end, *readability formulas* were originally created to predict the reading difficulty associated with text. All in all, readability study is concerned with ensuring that a given piece of writing reaches and affects its audience in the way that the author intends [4].

1.2 Basic methods for testing reading

Two areas of applied linguistic theories – reading and testing- come together when testers design a test of reading ability. In such cases, the test designer decides what s/he wants to test i.e. what s/he means by reading ability and finds a means of testing it. Alderson [5] points out that there is no ‘best method’ for testing reading comprehension and no single test method fulfill all the purposes of tests. Of significant methods of testing comprehension, one can refer to discrete-point (multiple-choice) and integrative (cloze) tests.

Multiple-choice method: multiple-choice questions are a common device for testing students’ reading comprehension. Despite of their popularity, the value and validity of multiple-choice method is under question. Kobayashi [6] and Alderson [5] argue that despite these tests popularity as a test format for assessing reading comprehension in a second/foreign language, they have a significant drawback in that test takers can guess the right answer without fully understanding the reading passage, and thus test validity is questionable.

*Corresponding Author: Sorayya Mozaffarzadeh, Department of English Language and Literature, University of Tabriz, Iran. Email: soraya.mozafarzadeh@yahoo.com

1.3 Cloze test

The term 'cloze procedure' was first developed by Wilson Taylor in 1953 which seems to be a spelling corruption of the word "close". He explains that the term cloze derived from the Gestalt psychology concept of 'closure' [7]. It describes a tendency that humans have to complete a familiar but not-quite-finished pattern [8]. In the cloze procedure, the closures are created by deleting certain words from a passage. The examinee, then, is required to fill in the blanks with appropriate words on the basis of contextual clues provided in the passage [9].

One of the major test forms which emerged from the Spolsky's 'post modern' trend in language testing has been the cloze tests. Cloze test also makes use of Spolsky's idea of reduced redundancy [10]; that is, knowledge of language requires the ability to function even when there is reduced redundancy; that is "language learner presented with mutilated language can use his/her acquired competence to restore either the original text or an acceptable text" [2]. The cloze test reduces natural linguistic redundancies and require the examinee to rely upon organizational constrains to fill the blanks and infer meaning [11]. Spolsky [10] believes that the more developed competence the learner has, the more able he is in making use of the clues provided by the text to restore a greater number of missed items.

Taylor was the first to study cloze procedure, in 1953, for its effectiveness as an instrument for determining the readability level of the texts. In the 1960s, it served as a device of assessing reading comprehension in L1 and L2. During the 1970s, cloze tests began to be used as a measurement of overall L2 proficiency (Ahluwalia, 1992, cited in [8]). Today, cloze tests are widely used in some places (such as Iran) and as part of some large-scale language tests (such as TOEFL, IELTS). After Taylor's introduction of cloze procedure' different types of cloze tests were developed which the significant ones are traditional cloze and discourse cloze tests.

In traditional cloze testing or fixed-ratio method or standard cloze, every n^{th} ($n= 5-10$) word of a passage is removed and replaced by a standard-length blank space [7]. Usually, no word is omitted either in the first or the last sentence of the passage to provide the examinee with some context. Students are required to supply either the original word of the author or an appropriate equivalent in the blank space. This kind of deletion is called random deletion because it deletes every n^{th} word consistently, so that all classes and types of words have an equal chance of being deleted. It is believed that this type of deleting provide an actual sampling of real-life language [7].

Discourse cloze or rational cloze, a modification of cloze procedure was introduced by Bachman in 1985 to measure specific linguistic abilities in reading assessment for example grammatical features [12]. It involves deletion of special words from the passage, to include a development of sensitivity to the operation of lexical items in the discourse. In this type of cloze, a specific type of word is deleted according to a linguistic principle, such as nouns, verbs, adjectives, etc [8].

1.4 C-test

Early empirical investigations into cloze tests produced promising results which caused the popularity of the tests. However, other investigations by Klein-Braley, 1981 showed that the cloze tests are not the universal panacea and that there are, in fact, considerable problems with this type of tests [14]. In the light of the criticism leveled at the cloze test, a modification has been proposed by Klein-Braley and Raatz in 1982. The new testing procedure called *c-test* based on the tenets of cloze test without its deficiencies. The construction of c-test involves a number of short texts (usually five or six) to which the *rule of two* is applied. According to this rule, the second half of every second word in the text is deleted until the required number of mutilations is reached, leaving the first and last sentence of passage intact to provide enough contexts [14]. These tests like cloze procedures have been used as tests of language proficiency and rarely as tests of reading comprehension.

The present study aims to study the effectiveness of cloze test vs. C-test as measures of reading comprehension. So through this study, I try to answer the following research question:

1. Is there any difference between advanced subjects' performance on the C-test and their scores on the cloze test as measures of reading comprehension?

METHODOLOGY

2.1 subjects

This study was conducted in Iran Language Institute (ILI), Tabriz Branch. The subjects of the study were 27 adult female EFL advanced learners. The proficiency levels of the subjects were determined by placement tests of the institute. Most of them were from Tabriz and were speakers of Azeri. The only opportunity for learners to communicate in English is this formal classroom interaction. They have no or little opportunity for informal interaction outside the classroom. They have to speak English in the classroom, and they are not allowed to use Azeri or Persian in the classroom.

¹ . The term 'cloze procedure' and 'cloze test' are closely related. Sometimes they are used interchangeably. The difference is that cloze procedure is more general, demarking the use of an activity that follow 'cloze procedure' while a cloze test is a specific application of the cloze procedure to the testing situation [13].

2.2 Data collection procedure

2.2.1 Data collection instruments

The used instruments in this study were 2 tests. The first test was standard multiple cloze tests. The cloze test was served as a counterpart device to be compared with the second test i.e. c-test. The difficulty level of the tests was in accordance with the proficiency level of the subjects.

2.2.2 Test preparation

For preparing tests, different texts were studied and finally two which were appropriate were selected. In the selection of texts, it was tried to take into consideration the factors proposed by Day [3], esp. the readability, the culturally suitable, and appearance factors. The readability of selected texts was calculated by Flesch Readability Formula.

The standard multiple-choice cloze test was developed out of the passages taken from authentic sources (TOFEL) by using a 7th deletion random cloze test. The first and second sentences were left intact to yield what Oller and Jonz [15] call lead-in and lead-out, and the deletions began with the 7th word of the second sentence. The difficulty levels of the text used for constructing cloze test was 46.3. The constructed tests yielded 38 items. The prepared test was piloted two times among the learners similar to the samples. At the first time the test was piloted in the traditional form referred to as the free-response cloze test. The most frequent incorrect responses written in the pilot was used to construct the distractors of the multiple-choice cloze. The distractors have the same part of speech. In the second time piloting, the malfunctioned distractors recognized and replaced with suitable ones. Furthermore, the problems with the appearance of the test were obviated and the required time for test completion was estimated.

The reason of using multiple-choice cloze was that it could provide the possibility of objective scoring. Because it was intended to be compared to C-tests which had the exact method of scoring, the cloze tests should be scored in exact method; however, the use of exact method of scoring required the testee to provide the original words deleted from the passage which made the test extremely difficult and therefore frustrating.

The C-test like cloze test was constructed out of the passages extracted from authentic source with approximate difficulty level to the cloze test. The difficulty level of the used text was 41.4. The C-test was constructed by using the rule of two referred to them as traditional C-tests developed by Klein-Braley in 1982. That is, it was developed by deleting second half of every second word. If a word had only one letter it was ignored in counting the words and if a word had an odd number of letters, the larger half were removed. The first and last sentences of the texts were left intact and the deletion began from the second word of second sentence. The constructed C-test had 112 items. The test also was piloted to detect the probable problems with the tests like incorrect typing and also to estimate the length of the time needed to complete the tests.

2.2.3 Test administration

The final versions of tests were administered with one week interval. To control the observer effect, the tests were administered by the subjects own teacher and in their own class time. The allocated time for completing the tests was 30 minutes. Of course, we could extend it if more time was needed because our intention was covering of whole tests by the subjects.

2.2.4 Test scoring

The used scoring method for both of the cloze test and C-test is exact word method. This method was objective so that obtained scores are reliable. The multiple-cloze test was scored like usual multiple-choice items and each item had one point. In the C-test also each item had one point. In this study we decided to tolerate the spelling problems in the C-test which did not changed the meaning and part of speeches of the words. If these two happened the written word did not gain the mark.

3. FINDINGS AND DISCUSSION

3.1 descriptive Statistics

After scoring, for the ease of comparison all the scores were calculated out of 100. Descriptive statistics for both cloze test and C-test are represented in Table 1.

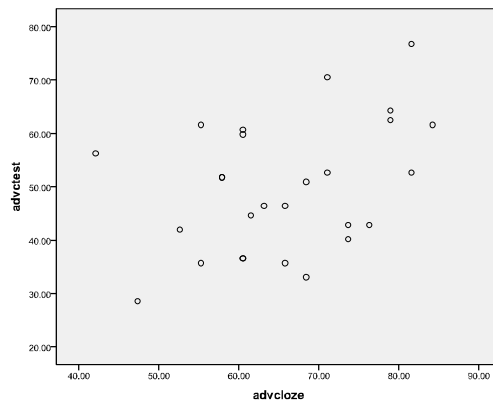
Table 1. descriptive statistics for subjects performance on the cloze test and C-test					
	N	Minimum	Maximum	Mean	Std. Deviation
Cloze test	27	42.10	84.21	65.7237	10.82079
C-test	27	28.57	76.78	49.8300	12.20135
Valid N (listwise)	27				

SD= Standard Deviation

As table one presents, the mean score of the subjects in cloze test (65.72) is higher than the mean score of them in the C-test (49.83). Table 2 shows the results of paired T-test and calculated correlation between cloze test and C-test.

		N	Correlation	Sig.
Pair 1	cloze & c-test	27	.384	.048

To investigate the degree of go togetherness of the two tests, a correlation analysis was performed of which the result is also shown in the chart below. The obtained correlation between the two tests is 0.38.



4. Conclusion and implications

The aim of this study was to answer the question that whether there is any difference between advanced subjects' performance on the C-test and their scores on the cloze test as measures of reading comprehension. The general conclusion and answer which can be drawn from the findings of this study is that despite the widely held view that c-test works better than the cloze test, it was shown that subjects performed better on the cloze test as measure of reading comprehension.

However, there it is required more investigation about differences of these test a long with the testees' opinion. Also further study should be done on C-test and cloze test as measures of reading comprehension with subjects with different levels of language proficiency.

REFERENCES

- [1] Lee, Y. (2004). Examining Passage-Related Local Items Dependence (LID) and Measurement Construct Using Q3 Statistics in an EFL Reading Comprehension test. *Language Testing*, 21(1), 74-100
- [2] Klein-Braley, C. (1985). A Cloze up on the C-test: a Study on the Constructive Validation of Authentic Tests. *Language Testing*, 2; 76
- [3] Day, R. R. (1994). Selecting a Passage for the EFL Reading Class. *Forum*, 32 (1). Retrived from <http://eca.state.gov/forum/vols/vol32/no1/p20.htm>.
- [4] Zamanian, M. and Heydari, P. (2012). Readability of Texts: State of the Art. *Theory and Practice in Language Studies*, 2(1), 43-53
- [5] Alderson, J. C. (2000). *Assessing Reading*. Cambridge,: Cambridge University Press.
- [6] Kobayashi, M. (2002). Methods Effect on Reading Comprehension Test Performance: Text Organization and Response Format. *Language testing*, 19 (2), 193-220
- [7] Oller, J. W. (1979). *Language Tests at School*. U.S.A.: Longman.
- [8] Lu, G. (2006). *Cloze Tests and Reading Strategies in English Language Teaching in China*. Thesis. University of the Western Cape, South Africa.
- [9] Farhady, H. (1986). Fundamental Concepts in Language testing (5) Theories. *Roshd Foreign Language teaching Journal*; 2(4), 28-37.

- [10] Spolsky, B. (1969). Reduced Redundancy as a Language Testing Tool. Presented at Second International Congress of Applied Linguistics, Cambridge, England, September 8-12.
- [11] Mousavi, S. A. (1999). *A Dictionary of Language Testing*. 2 ed., Tehran: Rahnama Publications.
- [12] Lee, S. H. (2008). Beyond Reading and Proficiency Assessment: the Rational Cloze procedure as Stimulus for Integrated Reading, Writing and Teacher-Student Interaction in ESL. *Elsevier (system)*, 36, 642-66
- [13] McMamey, T. (2006). Getting Closer on the Cloze: A Validity Study of the 'Rational Deletion' Method. *Second Language Studies*, 24(2), 114-164.
- [14] Klein-Braley, C. and Rattz, U. (1984). A Survey of Research on the C-test. *Language Testing*, 1(2); 134-146.
- [15] Oller, J. W., Jr., & Jonz, Jon. (1994). *Cloze and coherence*. Cranbury, New Jersey: Bucknell University Press.