Construct Validity of Multiple-Choice Cloze Test and Cloze-Elide Test in Testing Reading Comprehension among Iranian EFL Learners

Mostafa Saeedi

MA at TEFL, Ilam University, Iran

Corresponding Author email: Mostafasaeddy26@yahoo.com

ABSTRACT: Cloze tests have been extensively used for measuring readability, language proficiency and reading comprehension. This study is intended to determine and cross-compare the construct validity of multiple-choice cloze tests and cloze elide tests in testing reading comprehension via a differential group design. In this study, two instruments were developed and used: a multiple-choice cloze test and a cloze elide test. The tests were administered to sixty, BA and MA EFL Iranian students. The results indicated that a) multiple-choice cloze tests are valid in testing reading comprehension, b) cloze elide tests are valid in testing reading comprehension, c) BA and MA performances on cloze elide tests were better than BA and MA performances on multiple-choice cloze tests.

Keywords: Construct validity, Multiple-Choice cloze tests, Reading comprehension, Cloze elide test, Differential group design

INTRODUCTION

Testing is important in human lives, as well as in the language learning. Tests are used to make decisions which influence people's lives (Farhadi, Ja'fapour, Birjandi, 2006). Madsen (1983) says that tests can help students in at least two ways. First, creating positive attitudes toward classes and second English tests can help students to master the language. Language teachers also can benefit from language tests. Madsen (1983) argues that tests can help teachers to diagnose their own efforts as well as those of their students. As Heaton (1990) says tests can help teachers to diagnose the student's weaknesses and difficulties. In language learning, testing is important to find out a general idea about student's proficiency in the target language. It means that a test can be a tool to measure how successful a language teaching had been held.

Testing includes testing of four skills that is speaking, listening, reading, and writing. To test these skills several different techniques were developed. As in Iran reading gets more proportion of attention, so the focus of this study is on techniques related to testing reading comprehension.

Cloze Tests are common tests that are used for testing reading comprehension. And as tests are used for evaluative purposes, it is therefore essential to search about the validity of these tests in different contexts as much as possible. Decades of research have given us useful information about cloze tests; however, these issues in Iranian context are not informative enough and are opened to more research. So, this study focuses on two types of Cloze Tests, Multiple-Choice Cloze Test and Cloze-Elide Test, in testing reading comprehension skill.

Literature Review

Different definitions of the test mentioned by different researchers. Among these, Brown (2001) stated that a test is a method of measuring a person's ability or knowledge in a given domain. Method here refers to a set of techniques, procedures, and items that constitute an instrument of some sort that requires performance or activity on the part of the test-taker. In line with Brown, Heaton (1990) says tests are constructed primarily as devices to reinforce learning and to motivate the student or as a means of assessing the student's performance in the language. Based on these definitions, it can be inferred that testing is any activity of using tests to measure ability, knowledge, or performance.

As mentioned before, different types of test exist to test different language skills and Cloze tests are common tests that are used to test reading comprehension. In defining Cloze tests, Farhadi et al. (2006) defined...
cloze procedure as "a passage of appropriate difficulty (determined by readability formulas), and of appropriate length (220-250) with every seventh word deleted". There are, however, at least five main types of cloze tests available to language teacher: fixed-ratio, multiple-choice, rational, cloze elide, and C-test (Ikeguchi 1995; Weir 1990; Klein-Braley and Raatz 1984). But the main focus of this study is on multiple-choice cloze test and cloze elide test in testing reading comprehension.

The construction of multiple-choice cloze test begins with the construction of a normal cloze test. First, an appropriate passage is chosen. The deletion procedure begins with the second sentence. Every nth (5-10) word is deleted. The Second, stage requires the supplement of the deletion with three or more distracters. Finally, examinees will be required to choose the best answer that fits the context. Cloze elide tests also called "...intrusive word technique, text retrieval, text interruption, doctored text, and negative cloze'..." (Alderson, 2000). In this test, the tester inserts words and the test-taker is asked to find the words that do not belong to the text. In line with Alderson, Brown (2001) says that cloze-elide test is a kind of test that insert words to a text that actually do not belong to the text. The test-taker's task is to detect and cross out the intrusive words.

As said, since tests are used for decision making and these decisions affect test-takers' life so these tests need to be valid; that is, the test must actually measure what it is intended to measure. Validity also has different types: content validity, criterion-related validity, construct-related validity, and also face validity. This study concentrates on construct validity that is defined as the extent to which the items in a test reflect the essential aspects of the theory on which the test is based (i.e., the construct). Because the constructs occur in the brain and the test taker cannot directly take the constructs out of the students' brain the only way to reach constructs is indirect and through experiments. “That is the experiments are the most straightforward strategy available to testers for establishing construct validity” (Brown, 2005). The experiments take many forms but differential-groups, one of the easiest to understand, will be used in this study. Since even the best test can have damaging consequences if used inappropriately, in this respect test content is an important part of evaluating the appropriateness of a test for a given purpose. Therefore, the content validity of the tests were also determined through the judgment of experts. In sum, this study tries to investigate the construct validity of multiple-choice cloze tests and cloze elide tests in testing reading comprehension.

Many researchers have found that cloze tests can measure a wide range of language skills and competence. Reading requires learners to utilize clues available in language, and cloze procedure requires a similar ability (Rye, 1982). Alderson (2000) recommends cloze procedure for reading assessment. Studies by Alderson (2000), Yamashita (2003), Sadeghi (2010), Williams, Ari and Santamaria (2011), and Tabatabaei and Mirzaei (2014) showed that cloze tests have correlation with other reading test like TOFEL. And finally, Green (2001) claims that the findings of his study provide strong evidence that if cloze tests are designed appropriately, they permit valid assessment of reading comprehension. These are scholars who have supported cloze tests as a valid measure of reading comprehension primarily because of its correlation with other supposedly valid tests of reading comprehension.

There are, however, other studies that show contradictory findings about the suitability of cloze tests as a measure of high-level comprehension skills. For example, Lado (1986), who used a questionnaire as well as quantitative analysis of the results of cloze tests administered to college-level native speakers of English, found that cloze tests are inappropriate measures of native language proficiency and comprehension. Shanahan et al. (1982) by comparing intruded, natural, and scrambled cloze tests in three experiments, concluded that using cloze tests was unreasonable for global comprehension. Porter (1976) also states that “it would be misleading to suggest that the cloze procedure was measuring comprehension”.

As shown above, studies about cloze tests have controversial findings. Previous research also has more focused on concurrent validity of cloze tests. Because of these reasons, more studies need to be done about the validity of these tests in testing reading comprehension. That's why this study was design to determine the construct validity of a multiple-choice cloze test and a cloze elide test in testing reading comprehension through a differential group design.

Research questions
Are multiple-choice cloze tests valid for assessing test-takers’ reading comprehension ability?
Are cloze elide tests valid for assessing test-takers' reading comprehension ability?
Is there any significant difference between performances of the test-taker on the multiple-choice cloze test and cloze elide test?
METHOD

DESIGN

To answer the research questions guiding this study, a descriptive quantitative approach was developed. For this purpose, two tests, including a multiple-choice cloze test and a cloze elide test, were developed and administered to determine their validity via a differential group design. This study also sought to see whether there were any significant difference between performances of the test-taker on the multiple-choice cloze test and cloze elide test.

Participants

The subjects of this study were two groups of EFL Iranian learners of Shahid Chamran University at Ahvaz: the first year students of bachelor of arts (BA) and the last year students of master of arts (MA). All participants were native speakers of Persian. These two groups consisted of both male and female EFL learners. The age of the BA students were between 19 to 22 years old and the age of the MA students were between 25 to 33 years old. Each group consisted of 30 EFL learners.

Instruments

Two instruments were developed and used for this study: a multiple-choice cloze test and a cloze elide test. The passage of these two exams were selected from a general English textbook. After the selection of the texts almost every 7th word of each text were deleted, which created a text with 15 slots. Among these deletions numbers and special names were not counted. It is also worth mentioning that the first and the last sentence of each exam were left intact. Then the exams with deletions were given to EFL learners at Izeh Azad University to fill in the blanks with correct words. After that, the most frequent wrong answers were selected for distractors and intrusive words. Finally four choices, one correct answer and three distractors, for multiple-choice cloze test and two intrusive words, one correct and one incorrect, for cloze elide test were selected.

Procedure

Before the main administration, participants were explained in detail about the purpose of the study and the importance of their response. The tests were administered in the regular classrooms of the students during regular class hours. After the explanation, each group received two exams, multiple-choice cloze test and cloze elide test. Each group had 30 minutes to do the tests, to each test 15 minutes allocated. After the distribution, the researcher reminded the participants that the wrong answers will not receive negative score.

Data analysis

The gathered data through multiple-choice cloze test and cloze elide test analyzed by the statistical package for social sciences (SPSS), 21th version of this software used in this respect. An independent sample T-test were used to compare means between the performances of these two groups on multiple-choice cloze test and cloze elide test and also to see whether there is a significant difference between performances of these two groups on multiple-choice cloze test and cloze elide test. A paired sample T-test were also used to compare the performances of each group on both tests.

RESULTS

Table 1 represents the descriptive statistics of MA and BA performances on Multiple-Choice Cloze Test.

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>Variance</th>
<th>Std. Error</th>
<th>Maximum</th>
<th>Minimum</th>
</tr>
</thead>
<tbody>
<tr>
<td>MA</td>
<td>30</td>
<td>8.167</td>
<td>2.81723</td>
<td>7.937</td>
<td>.51435</td>
<td>13</td>
<td>4</td>
</tr>
<tr>
<td>BA</td>
<td>30</td>
<td>5.133</td>
<td>1.71672</td>
<td>2.947</td>
<td>.31343</td>
<td>9</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 1 shows that the mean of MA performances on Multiple-Choice Cloze Test (M= 8.1667) is bigger than the mean of BA performances on Multiple-Choice Cloze Test (M= 5.1333). From the above results it can be concluded that the MA performances on Multiple-Choice Cloze Test is better than the BA performances on
Multiple-Choice Cloze Test. But these results are not sufficient to answer the first research question. To do so an Independent T-Test was run.

The results of independent sample t-test, comparing the mean score of the MA (N=30, M=8.16, SD=2.81) and the BA (N=30, M=5.13, SD=1.71) students’ performances on Multiple-Choice Cloze Test show that the Sig (2-tailed) is equal to .000 and degree of freedom (df), and t are equal to 47.927 and -5.036 respectively. As the Sig (2-tailed) is less than .05, it can be said that the difference between mean of the MA and the BA students’ performances on Multiple-Choice Cloze Test is significant. That is MA performances on multiple-choice cloze tests are significantly different from BA performances on these tests.

Table 2 presents the descriptive statistics of MA and BA performances on cloze-elide test, related to the second research question.

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>Variance</th>
<th>Std. Error Mean</th>
<th>Maximum</th>
<th>Minimum</th>
</tr>
</thead>
<tbody>
<tr>
<td>MA</td>
<td>30</td>
<td>10.2667</td>
<td>2.76597</td>
<td>7.651</td>
<td>.50499</td>
<td>14</td>
<td>3</td>
</tr>
<tr>
<td>BA</td>
<td>30</td>
<td>7.7667</td>
<td>2.07918</td>
<td>4.323</td>
<td>.3796</td>
<td>12</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 2 indicates that 60 participants, 30 MA and 30 BA, took part in cloze-elide test. The mean of MA performances on cloze-elide test is 10.2667 (M=10.2667) and the standard deviation and the variance are 2.76597 and 7.651 respectively. The mean score of BA performances on cloze-elide test is 7.7667 and the standard deviation and the variance are 2.07918 and 4.323 respectively. To make comparison between the mean score of BA and MA performances on cloze-elide test, an independent t-test was run.

As indicated in Table 2 the mean of MA performances on cloze-elide test (N=30, M=10.26, SD=2.76) was greater than the mean of BA performances on cloze-elide test (N=30, M=7.76, SD=7.07). Since the Sig (2-tailed) in Table 4.11 is .000 and less than .05 (.000<.05) it can be inferred that these differences are significant, and from the above results it can be said that the differences between MA and BA performances on cloze elide tests are significantly different.

"Is there any significant difference between performances of the test-takers on the multiple-choice cloze test and cloze elide test?" is the third research question of this study. To answer this research question first the performances of MA students on multiple-choice cloze test compared to their performances on cloze elide test and then the same procedure was done for BA students that is the performances of BA students on multiple-choice cloze test compared to their performances on cloze elide test.

Table 3 indicates the descriptive statistics of MA performances on both multiple-choice cloze test and cloze elide test.

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>Variance</th>
<th>Std. Error Mean</th>
<th>Maximum</th>
<th>Minimum</th>
</tr>
</thead>
<tbody>
<tr>
<td>MA</td>
<td>30</td>
<td>8.1667</td>
<td>2.81723</td>
<td>7.937</td>
<td>.51435</td>
<td>13</td>
<td>4</td>
</tr>
<tr>
<td>BA</td>
<td>30</td>
<td>10.2667</td>
<td>2.76597</td>
<td>7.651</td>
<td>.50499</td>
<td>12</td>
<td>3</td>
</tr>
</tbody>
</table>

As it is clear in Table 3, the mean score and the standard deviation of MA performances on Multiple-Choice Cloze Test are 8.1667 and 2.81723. and on Cloze-Elide Test, the mean score and the standard deviation of MA performances, are 10.2667 and 2.76597. From these results it can be deduced that the MA performances on Cloze Eilde Test is better than their performances on multiple-choice cloze test. A paired t-test was run on a sample of 30 MA students to compare their performances on multiple-choice cloze test and cloze elide test.

The results of paired-sample t-test of the MA performances on cloze elide test and on multiple-choice cloze Test. The degree of freedom (df) and the statistical t (t) are 29 and -2.768, and the Sig (2-tailed) is .010. Since the Sig (2-tailed) is less than .05 (p-value < .05), it can be said that there is a significant difference between MA performances on cloze-elide test and their performances on multiple-choice cloze test.

Table 4 represents the descriptive statistics of BA performances on multiple-choice cloze test and on cloze elide test.

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>Variance</th>
<th>Std. Error Mean</th>
<th>Maximum</th>
<th>Minimum</th>
</tr>
</thead>
<tbody>
<tr>
<td>MA</td>
<td>30</td>
<td>5.1333</td>
<td>1.71672</td>
<td>2.947</td>
<td>.31343</td>
<td>9</td>
<td>2</td>
</tr>
<tr>
<td>BA</td>
<td>30</td>
<td>7.7667</td>
<td>2.07918</td>
<td>4.323</td>
<td>.37960</td>
<td>12</td>
<td>2</td>
</tr>
</tbody>
</table>
As shown in Table 4 the mean score and the standard deviation of BA performances on multiple-choice cloze test are 5.1333 and 1.71672. For BA performances on cloze-elide test the mean score and the standard deviation are 7.7667 and 2.07918. From these results it is clear that the mean score of BA performances on cloze-elide test is greater than the mean score of BA performances on multiple-choice cloze test, in other words the BA performances on cloze-elide test and on multiple-choice cloze test are different. But to see whether this difference is significant or not a paired t-test was run.

The results of paired t-test indicated that the statistical t (t) and the degree of freedom (df) are -5.111 and 29, and the Sig (2-tailed) is .000. Since the Sig (2-tailed) is less than .05 (p-value < .05), it can be said that there is a significant difference between BA performances on cloze-elide test and their performances on multiple-choice cloze test.

**DISCUSSION**

This study was, in fact, an attempt to investigate the construct validity of two varieties of cloze test; that is, multiple choice cloze test and cloze-elide test in testing reading comprehension.

Based on the results of this study there was a significant difference between MA and BA performances on multiple-choice cloze test. So it can be concluded that multiple-choice cloze tests are valid tests in testing reading comprehension. The results, also, made it clear that the differences between the performances of MA and BA students on cloze elide test were significant this means that cloze elide tests are valid in testing reading comprehension.

As it was proved that both multiple-choice cloze test and cloze elide test were valid in testing reading comprehension, the performances of MA and BA students on these two tests were compared.

The results indicated that the performances of MA and BA students on cloze elide test were better than their performances on multiple-choice cloze test.

Though there where different and controversial findings about the validity of cloze tests in testing reading comprehension, the analysis of the data of this study made it clear that multiple choice cloze test and cloze-elide test are valid tests in testing reading comprehension.

The findings of this study support the previous findings of (Scholars, 1987; Rye, 1982; Alderson, 2000; Yamashita, 2003; Sadeghi, 2010; Williams et al., 2011; Tabatabaei and Mirzaei, 2014; and Green, 2001) stating that cloze tests are valid tests in testing reading comprehension though these studies mainly used correlational methods.

However these findings are in contrast with the findings of (Shanahan et al. 1982; Lado, 1986; Porter, 1976) asserting that cloze tests are not valid tests in testing reading comprehension.

Such findings in language testing research can be interpreted as these two tests can be used in testing reading comprehension. The preceding statistical data would, therefore, mean that multiple cloze test and cloze elide test are valid measures of EFL reading comprehension.

It is worth mentioning that the findings related to third research question showed that MA and BA performances on multiple choice cloze test and cloze elide test are not the same. MA and BA performances on cloze elide test was better than their performances on multiple-choice cloze test. This may be due to the fact that answering a question with two options and deleting the wrong one in cloze elide test is simpler than choosing the correct one among four options.

**CONCLUSION**

The aim of this study was to answer the question of the validity of multiple-choice cloze test and cloze elide test in testing reading comprehension. Based on the analysis of the data the researcher draw some conclusion: first, the result of MA and BA students' test on multiple-choice cloze test and cloze elide test shows that these two tests are valid in testing reading comprehension. Second, MA performances on multiple-choice cloze test and cloze elide test was compared. The result shows that there is a significant difference between MA performances on cloze-elide test and on multiple-choice cloze test. Third, BA performances on multiple-choice cloze test and cloze elide test was compared. The result shows that there is a significant difference between BA performances on cloze-elide test and on multiple-choice cloze test. The result of the last two conclusion is that the MA and BA performances on cloze elide test was better than the MA and BA performances on multiple-choice cloze test. In sum these results led the researcher to conclude that multiple-choice cloze test and cloze elide test are valid in testing reading comprehension.
On the basis of the research findings and discussion it was concluded that multiple-choice cloze test and cloze elide test are valid in testing reading comprehension, so teachers can safely use these two tests in testing reading comprehension. Also teachers can use these two tests in determining students’ weaknesses and strengths in reading comprehension.

The researchers recommended that this study can be expanded on larger populations, and as cloze tests can be presented variously, especially the way of presenting the options of the answer, so the teachers can adapt based on their student’s condition.

REFERENCES