

Abstract

During the past fifty years, teacher education has been the focus of many studies and has drawn the attention of many researchers. As a result, two main conceptions of teacher education have emerged in the field. One approach which is referred to as training emphasizes the traditional views to teacher education. The second view known as development considers the teacher as an active processor of his/her own teaching practices through self-awareness and reflection. In order to better grasp the concept of these two perspectives, as well as their differences, this paper outlines and contrasts the two approaches with a focus on the role of the teacher in each approach. Then, the drawbacks of the *training* perspective are discussed with the aim of drawing the attention of teachers and educational authorities to the benefits of the developmental view of teacher education. Finally certain implications of the *development* perspective are discussed in relation to the process of language teaching and learning both for teachers and for those who are involved in the planning of teacher education programs.

Key Words: language teacher, teacher education, teacher development, teacher training

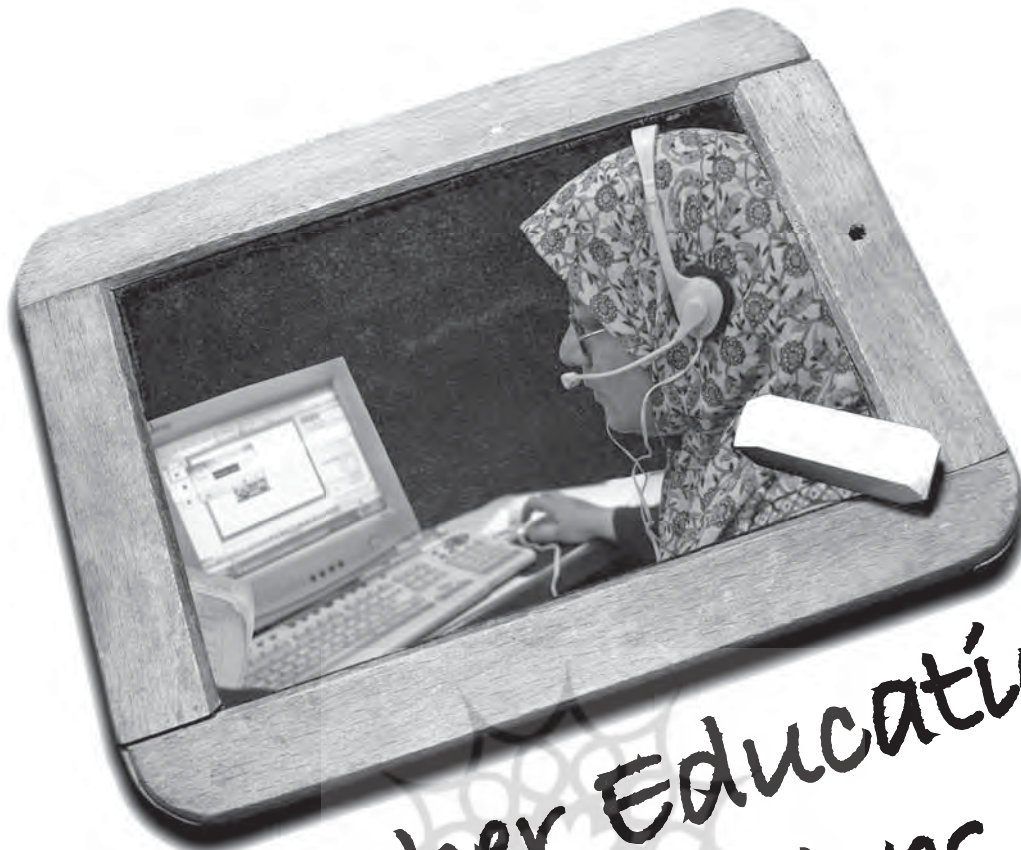
Introduction

Nowadays, education is seen as the basis of every progress in human societies which requires special attention by the authorities in order to provide an appropriate educational system capable of meeting the educational needs of people. Education can be considered interaction between students as the recipients of education, and, the teachers as providers of education. In fact, the success of any educational system particularly in terms of quality depends, to a large extent, on the provision of qualified teachers. No nation can achieve economic, social, technological, and scientific progress without a good system of education to sustain its achievement. Clearly, in this direction the main burden is on the teachers who have the responsibility of providing high quality instruction. However, this is not an easy task to achieve.

With technological advancements, the

increasing use of electronic sources of information and emergence of new fields of study, the need for more and better teachers is felt more than ever before. Also, the need to raise the level of knowledge and skills of the existing teachers is seriously felt. Teachers, therefore, need more opportunities than ever before to continue learning throughout their careers. In fact, teacher education should be taken into account both by the authorities, and, especially, the teachers themselves; certain instructional programs should be designed in order to enhance the quality of teaching, as well as to make teachers function better in their classrooms.

According to Freeman and Johnson (1998), teacher education refers to the policies and procedures designed to equip prospective teachers with the knowledge, attitudes, behaviors, and skills they require to perform more effectively in the classroom, school, and wider community.



Teacher Education Perspectives

پژوهشگاه علوم انسانی و مطالعات فرهنگی
رتال جامع علوم انسانی

Parviz Maftoon PhD in TEFL, Islamic Azad University,
Science & Research Campus
Pmaftoon@gmail.com
Seyyede Najafi Sarem



چکیده

طی پنجاه سال گذشته، موضوع تربیت معلم، همواره مورد توجه بسیاری از محققان و متخصصین بوده است. در نتیجه این مطالعات، دو مفهوم اصلی در زمینه تربیت معلم پدید آمده است. اولین مفهوم، رویکردی است که به آموزش معلم از نظر سنتی تأکید دارد و دومین مفهوم به رویکردی می‌پردازد که معلم را به عنوان یک پردازشگر فعال از طریق تأمل و خودآگاهی می‌داند. در این مقاله، سعی شده است برای درک بهتر این دو دیدگاه، تفاوت‌ها و تمایزهایشان با توجه به نقش معلم در هر یک از این دو دیدگاه بیان شوند.

نقاط ضعف و قوت آموزش معلم، با هدف جلب توجه معلمان و متخصصان آموزش به بحث گذاشته می‌شود و در آخر توصیه‌هایی در مورد رشد معلم در ارتباط با فرایند یادگیری و آموزش مورد بررسی قرار می‌گیرد.

کلیدواژه‌ها: مدرسان زبان، تربیت معلم، رشد معلم، آموزش معلم

a. It is recommended that cloze tests be introduced in all English classes even at elementary level. English teachers should include different varieties of cloze test in classroom instruction and testing sessions.

b. For preparing of the cloze tests, the inclusion of materials drawn from familiar topics for which students have presuppositions and schemata (Schema familiarity), particularly, those related to general English should be useful.

c. Cloze tests alone cannot give a true picture of all language skills ability of EFL learners. Cloze test irrespective of its format should be accompanied with other types of English tests.

d. The readability level of passages along with the degree of cohesiveness, language variety (spoken versus written), and text difficulty should be taken into consideration in the preparation process or designing any cloze test.

Reference

- Anderson, J. (1971). A technique for measuring reading comprehension and readability. *English Language Teaching Journal*, 25(2), 178-82.
- Bachman, L.F. (1985). Performance on cloze tests with fixed-ratio and rational deletions. *TESOL Quarterly* 19(3): 535-555.
- Bachman, L.F. (2000). Modern language testing at the turn of the century: assuring that what we count counts. *Language Testing* 17, 1-42.
- Bachman, L.F. (2004). *Statistical analyses for language assessment*. Cambridge: Cambridge University Press.
- Brown, J. D., Yamashiro, A. D., & Ogane, E. (2001). The emperor's new cloze: Strategies for revising cloze tests. In T. Hudson & J. D. Brown (Eds.), *A focus on Language test development: Expanding the language proficiency construct across a variety of tests*. 89-118. Honolulu, HI: University of Hawaii Press.
- Brown, J. D. (2002). Do cloze tests work? Or is it an illusion? *Second Language Studied*. 21 (1) 79-125.
- Brown, J., & Eskenazi, M. (2004). Retrieval of Authentic Documents for Reader-Specific Lexical Practice. In *STIL/ICALL Symposium* (Vol. 2). Venice, Italy.
- Heaton, J.B. (1990). *Writing English Language Tests*. Rowley, MA: Newbury House Publishers.
- Higgins, D. (2006). *Item Distiller: Text retrieval for Computer-Assisted Test Item Creation*. ETS, Princeton, NJ.
- Kobayashi, M. (2002). Cloze tests revisited: Exploring item characteristics with special attention to scoring methods. *Modern Language Journal*, 86(4), 571-586.
- Kunnan, A. J. (2000). *Fairness and Validation in Language Assessment: Selected papers Foreign Language Study*.
- Liu, C., Wang, C., Gao, Z., & Huang, S. (2005). Applications of Lexical Information for Algorithmically Composing Multiple-Choice Cloze Items. In *Proceedings of the 2nd Workshop on Building Educational Applications Using NLP* (p. 1-8). Ann Arbor, MI: Association for Computational Linguistics.
- Mirhassani, A & Reshadi, Z. (2001) the relationship between Genre and Iranian EFL learners' performance on open-ended and multiple-choice cloze tests. *IJAL*.5, (1).68-82.
- Nakagawa, T. (2002). Interaction between the cloze test and expectancy grammar of Japanese EFL learners. *Research bulletin of Tsukuba International University*. Bulletin of Tsukuba International University.
- Oller, J.W. Jr., & Conrad, C. (1971). The cloze technique and ESL proficiency. *Language Learning* 21, 183-195.
- Oller, J. W., Jr., & Jonz, Jon. (1994). *Cloze and Coherence*. Cranbury, New Jersey: Bucknell University Press.
- Spolsky, B. (1978). Introduction to linguists and language testers. *Papers in Applied Linguistics*, Advances in language testing.
- Steinman, L. (2002). Considering the cloze. *Canadian Modern Language Review* 59(2). Retrieved January 20, 2006,. From http://www.utpjournals.com/product/cmlr/592/592_TCL_Steinman



was supported. It means that manipulating the blank length and even presenting the initial letter of each blank had no significant impact on the reliability and validity of the cloze test. It can be concluded that these four versions are similar with regard to their reliability and validity; hence, none of them are highly superior to the others. Therefore, it is safe to claim that the above results indicate that the cloze test we have developed for assessing the proficiency of Iranian foreign language learners of English meets both testing and L2 research standards. First, the cloze test appears to be valid. Comparison of the cloze test scores with the standardized proficiency test scores provided evidence for the validity of the cloze test. Moreover, the cloze test is reliable: its overall internal consistency index is quite high, and it remained considerably high even when blank length and formats were manipulated, suggesting that the cloze test could also be used to assess the proficiency of Iranian EFL learners in a single classroom level. Furthermore, the cloze test has the capacity to discriminate between EFL learners in a wide

range of proficiency levels, yielding a wide distribution of scores both by participant and by item (Kunnan, A. J. 2000). Finally, the cloze test is practical, requiring EFL learners less than 50 minutes to complete. All in all, the present cloze test appears to be an excellent tool for assessing the proficiency of college-level Iranian EFL learners.

Pedagogical Implications

The results of this study indicated that there is no relationship between blank length manipulation and reliability and validity of cloze test. That is, the difference between the reliability and validity indices of the four forms of cloze tests was not significant. Consequently, EFL/ESL teachers can use all these forms in their classes. For the sake of variety, teachers can use them and be sure that these different versions of the same test all enjoy satisfactory reliability and validity. However, concerning practicality, the test form A is more practical than the other forms because in this form all blanks are of uniform length. Providing standard blanks is simpler for Iranian English teachers than providing other varieties of blanks. It takes less effort and time. The implication is that trainee teachers should develop and even manipulate different standard cloze tests for measuring the progress of their students even in classroom settings.

Recommendations for English teachers in Iran using cloze tests

The following recommendations can be made based on the results of the present study below:

**A
cloze test
is considered as an
integrative test because
it includes the context and
sociolinguistic rules of that
particular language while
assessing the language
ability of the learners**

none of the tests was different from the other tests in terms of reliability.

Table 3 Cloze tests Validity with Michigan Language Proficiency Test

Test	Reliability
Form A	0.58
Form B	0.58
Form C	0.58
Form D	0.70

To compare the differences among the correlation coefficients of tests, the Fisher's transformation to Z formula (Guilford, 1978) was used. The following table presents the validity indices and their corresponding Z coefficients.

Table 4 Conversion of R into Z coefficient

Test	R	Z
Form A	0.58	0.66
Form B	0.58	0.67
Form C	0.58	0.67
Form D	0.70	0.86

To test the significance of the difference between the two independent correlations each pair of Computed Z's from the respective correlation coefficients were compared. On the whole, six comparisons were made.

Table 5 Correlation coefficients comparison

A&B	0.06	Z critical	1.64
A&C	0.07		
A&D	1.29		
B&C	0.01		
B&D	1.26		
C&D	1.19		

The results revealed that there was no significant difference among the validity indices. That is, none of the correlation coefficients exceeded the Z critical, which was 1.64 at the .05 level of significance. Then, from the results of this study, it can be claimed that the blank length manipulation and presenting cues at the beginning of the cloze blanks do not affect the psychometric characteristics of cloze tests. Therefore, the null hypothesis stating that there is no relationship between the blank length manipulation and the reliability and validity of cloze tests was confirmed. Hence, it can be concluded that these four formats are similar with regard to their reliability and validity.

Conclusions

The results revealed that there was no significant difference among the reliability and validity indices of the four versions of the cloze test used in this study. It is quite logical to claim that the hypothesis, stating that there is no relationship between the blank length manipulation and test qualities,

Cloze test, which combines the advantages of integrative testing and objective scoring, was used as a device for estimating the readability level of written materials for native speakers of English

tests in the following order, first the cloze test and then the Michigan test. Participants were 45 junior and 35 senior students majoring in English. A total of 120 minutes was allocated to answering the combined 40 items of the cloze and 100 items of the Michigan test (25 minutes for the cloze test and 95 minutes for the Michigan test). In this study, the exact answer scoring method was used for the cloze, which meant that only the word that had occupied the blank in the original passage was numbered as correct. In scoring the Michigan test which consisted of 100 multiple choice items only the correct answers were counted. The participants were not penalized for their wrong responses, but the missed items were numbered as wrong. The next step of the test validation was to compute the psychometric characteristics of the cloze test. It showed the reliability of .65 and the validity of .60. The obtained values revealed that the test was satisfactory enough to be used in the present study.

The four versions of the cloze test were developed and were labeled as A, B, C, and D. After preparing the four forms of the cloze test, they were duplicated and randomly ordered so that all the participants would have an equal chance of getting any one of the four forms. Participants who took the test were 160 juniors and 160 seniors majoring in English. On the basis of the four formats of the cloze tests, the participants were assigned to four groups. Each format of the cloze test was answered by 80 participants. All the participants took the two tests, cloze and Michigan, concurrently.

Data Analysis

To test the null hypothesis, the reliability and validity indices of the four forms of the cloze test were estimated and compared. Therefore, the reliability of all tests was computed through Kuder-Richardson formula (KR-21) (table 1).

Table 1 Cloze Tests Reliability

Test	Reliability
Form A	0.61
Form B	0.60
Form C	0.66
Form D	0.58

Moreover, they were validated against the Michigan language proficiency test. The results are presented in the following tables. Thus using the F ratio (Guilford, 1978), the reliability indices of the tests were compared two by two. There were four cloze tests in this study; therefore, 6 comparisons were made.

Table 2 F ratio comparison of reliability indices

A&B	$F = 18.83/16.72 = 1.12$	F critical 2.60
A&C	$F = 21.16/16.72 = 1.26$	
A&D	$F = 20.70/16.72 = 1.23$	
B&C	$F = 21.16/18.83 = 1.12$	
B&D	$F = 20.70/18.83 = 1.09$	
C&D	$F = 21.16/20.70 = 1.02$	

First, reliability indices of tests A and B were compared. The obtained F value did not exceed the F critical, which was 2.60 at the .05 level of probability. Comparing the reliability indices of tests A and C and D and computing the F value revealed that

in this version an initial letter of the missing word was provided along with seven spaces. In the third and fourth versions there was a relationship between the number of letters of the missing words and the length of the blank. The length of the blank was equal to the number of letters of the missing word. The fourth version was quite similar to the third except that the initial letters of the missing words were provided.

Participants

The participants in this study were 320 Iranian students studying English at Kermanshah Islamic Azad University in Iran. All these students had already taken six to seven years of formal English courses prior to entering the university. Moreover, the participants were selected from among the English majors who had already passed at least 30 English courses at the university. They were both male and female junior and senior students majoring in English. The sample was homogeneous with regard to nationality, native language background, and field of study.

Instrument

Two tests were used in this study. One was The Michigan Test of English Language Proficiency, form "A" published in 1961, and the other was a cloze test having four versions.

Procedure

The procedure can be divided into two sections: procedure of test preparation and procedure of test administration.

a) Test preparation

To develop the cloze test, it was necessary to identify the appropriate difficulty level of the cloze passage. Several steps were taken to identify the appropriate difficulty level of the cloze passage. These steps are presented below.

1. Most of the books studied by juniors and seniors majoring in English were identified.

2. Ten paragraphs having at least 100 words were chosen randomly from each book.

3. The readability of the paragraphs was estimated through Fog index.

4. The average readability of each book was computed. That is, the readability of each individual book was identified.

5. The average readability of all books was computed.

6. The last step, a passage having the same readability level was selected.

A seventh word deletion pattern was used to create a cloze test from the passage. The cloze had a total of 40 items. The first two sentences and the last sentence were left intact to provide more context. Then, directions were added to tell the participants what to do in taking the cloze test.

b) Test Administration

Test administration was done in two different stages of test validation and administration of the final version of the cloze test. In test validation, the newly developed cloze test with uniform blank length was validated against the Michigan test. These two tests were administered concurrently. The participants took the

Since its emergence, the cloze procedure has been used for a multitude of purposes. Cloze test, which combines the advantages of integrative testing and objective scoring, was used as a device for estimating the readability level of written materials for native speakers of English. Extending this idea, Anderson (1971) used the cloze procedure to determine the readability of materials for an ESL class or individual. In the 1960s and 1970s, yet another line of research developed which investigated the effectiveness of cloze as a test of overall language proficiency in a foreign language (Oller & Conrad 1971). Although there is some controversy as to exactly what aspects of linguistic competence cloze tests measure, one important finding of testing research is that cloze test scores tend to correlate highly with standardized proficiency scores (e.g. Bachman, 1985, 2000). Assuming that the latter indeed

cloze tests also provide a valid proficiency measure. Brown, Yamashiro, and Ogane, (2001) state that by manipulating the difficulty level of the test, it is also possible to tailor the test for L2 learners at a particular proficiency level.

In spite of the increasing optimism regarding the usefulness of the cloze procedure, a number of significant questions remain to be answered. For instance, the validity and reliability of the cloze test is a controversial issue which many applied linguists and foreign language researchers (Oller, 1972b; Bachman, 1985; Kobayashi, 2002) have concerned themselves about for many years.

Blank length manipulation is considered as a potential variable influencing cloze test quality. Although the number of studies on cloze tests is numerous, no research regarding the possible effect of the blank length manipulation on validity and reliability of cloze tests has been reported so far. Following the shared goal of producing more valid and reliable cloze tests, the present study aimed to explore the effect of blank length manipulation on the reliability and validity of cloze tests.

The term 'cloze' stems from the notion of 'closure' in Gestalt psychology (Nakagawa, Takeshi. 2002). It refers to the human tendency to see and imagine every thing in its complete form

Method

In this study four versions of cloze test were developed. The main differences among these four versions were the length of the blanks and initial clues. The first version of the cloze test had uniform blanks that is to say the length of all blanks was limited to seven spaces. The same length was applied to the second version; however,

tap L2 proficiency (an assumption which, admittedly, can be questioned for oral skills), this correlation then suggests that

Abstract

The main focus of this study was to investigate if cloze tests in which the blanks corresponding to the length of the missing words or blanks with initial letters would result in more valid and reliable cloze tests. The results of the analysis indicated that manipulating the blank length and even presenting the initial letter of each blank had no significant impact on the reliability and validity of cloze tests. There was also evidence that showed there were significant differences in performance of the groups of participants on different forms of cloze tests with blank length variation.

Key Words: blank length variations, reliability, validity, cloze tests, language testing

Introduction

Teaching and testing are so closely interrelated that, as Heaton (1990) believes, it is impossible to work in either field without being concerned with the other. Language testing tends to follow the trends in language teaching. With the emergence and the development of new methods of language teaching, new testing methods have appeared. As Spolsky (1978) states, testing history can be divided into three major eras: the pre-scientific era, the psychometric structuralist era, and the integrative sociolinguistic era. These trends or approaches tried to develop tests which could assess the real ability of language learners adequately.

A cloze test is considered as an integrative test because it includes the context and sociolinguistic rules of that particular language while assessing the language ability of the learners. Wilson Taylor first coined the term “cloze” in 1953. He used this term, which was an intentional misspelling of the word “close”, and called this technique a new psychological

tool for measuring the effectiveness of communication (Steinman, L. 2002). A cloze procedure consists of a passage from which every ‘*nth*’ word is deleted. In the first two or three sentences, no deletions are made to allow the theme of the passage to become established. Then words are deleted from the passage based on the length of the intervals. The sentence from which a word is deleted can be referred to as the ‘stem’, and the removed term itself as the ‘key’ (Higgins, 2006). The portion of the sentence from which the key has been removed is the ‘*blank*’. The testees’ task is to fill in the blanks with appropriate terms. The term ‘*cloze*’ stems from the notion of ‘*closure*’ in Gestalt psychology (Nakagawa, 2002). It refers to the human tendency to see and imagine every thing in its complete form. Following this principle, in a cloze test, the testee is exposed to an incomplete text, that is, a text with a number of blanks. And then s/he is supposed to use his/her linguistic ability to guess the missing words and complete the text (Mirhassani & Reshadi, 2001).



The Effect of Blank Length Manipulation On the Reliability and Validity of Cloze Tests



Researchers

Kamran Janfeshan PhD in TEFL
Islamic Azad University, Kermanshah Branch
K_janfeshan@yahoo.com

پژوهشگاه علوم انسانی و مطالعات فرهنگی

چکیده

ارزش‌یابی زبان، هم‌بستگی بسیار بالایی با آموزش زبان دارد. به‌طوری‌که با به‌وجود آمدن شیوه‌های جدید تدریس زبان، بالطبع روش‌های جدید ارزش‌یابی هم متحول می‌شوند. هدف اصلی این تحقیق بررسی آزمون‌های بسته‌ای است که از ابتدای پیدایش به منظور تحقق اهداف متفاوت به کار گرفته شده‌اند. علی‌رغم خوش‌بینی فزاینده در مورد مفید بودن آزمون‌های بسته، موضوع روایی و اعتبار آن‌ها توجه محققان زبان‌های خارجی را برای سال‌های متمادی به خود جلب کرده است. پیدا کردن عوامل تأثیرگذار بر عملکرد آزمون‌دهندگان در آزمون‌های بسته، هدف مشترک بسیاری از این مطالعات است.

تحقیق حاضر کوشیده است رابطه بین تغییرات طول جاهای خالی آزمون‌های بسته و اعتبار و روایی آن‌ها را بررسی کند. بدین منظور، پس از تهیه یک متن با درجه خواندن مناسب، یک آزمون بسته با چهار شکل متفاوت ساخته شد. در فرمت (الف)، طول جاهای خالی یکسان، در شکل (ب)، طول جاهای خالی یکسان و اولین حرف کلمه افتاده ارائه شده بود. در فرمت (ج) تعداد حروف کلمات افتاده مشخص شده بود و در فرمت (د) تعداد حروف و اولین حرف کلمه افتاده ارائه شده بود. نتایج نشان داد که تغییرات اندازه طول جاهای خالی و ارائه حرف اول کلمات افتاده، تأثیری در روایی و اعتبار آزمون بسته ندارد. علاوه بر این، عملکرد زبان‌آموزان در شکل (د) بسیار چشم‌گیرتر از دیگر فرمت‌ها بوده است. بنابراین، معلمان زبان انگلیسی می‌توانند با اطمینان خاطر از درجه روایی و اعتبار، از شکل‌های متفاوت آزمون‌های بسته استفاده نمایند.

کلیدواژه‌ها: تغییرات طول جاهای خالی، روایی، اعتبار، آزمون‌های بسته، آزمون زبان