

RESEARCH

Open Access



Exploration of vocational high school students experiencing difficulty in cloze test performances: a mixed-methods study in Taiwan

Kuo-Zheng Feng^{1*}

*Correspondence:
jack03150402@gmail.com

¹Department of English,
National Chengchi University,
Taipei, Taiwan

Abstract

This study addressed a gap in existing research on Multiple-Choice (MC) cloze tests by focusing on the learners' perspective, specifically examining the difficulties faced by vocational high school students (VHSs). A nationwide sample of 293 VHSs participated, providing both quantitative and qualitative data through a self-developed questionnaire. The results revealed that vocabulary and grammar posed the greatest challenges in the MC cloze test, while sentence patterns were perceived as the least difficult by VHSs. Factors contributing to these difficulties included the need for increased focus on vocabulary and grammar learning. Some participants attributed challenges to personal perceptions of intellectual capability, while others highlighted the influential role of teachers' attitudes on their learning motivations and outcomes. The study suggested implications for test designs and teaching approaches. Despite these contributions, the study acknowledged limitations and offered suggestions for future research directions.

Keywords: MC cloze test, Learners' perspectives, Vocational high school students, Test takers' difficulty, Test designs

Introduction

Vast numbers of publications in English have produced a growing awareness of language learners' needs for comprehending various types of texts (Huttner, 2008). Constructing a high-stakes test with different types of reading passages could further strengthen this notion, and a positive backwash could be expected from those tests (Hughes & Hughes, 2020; Madsen, 1983). Most test takers in EFL contexts believe that the process and preparation for English learning should be largely directed toward the contents of those high-stakes tests (Kohonen, 1999), where reading comprehension is perceived to play a significant role (Lee, 2004). High-stakes tests in Taiwan, such as the General Scholastic Ability Test (GSAT) and the Technological and Vocational Education Joint College Entrance Examination (TVE-JCEE), are recognized as having an enormous influence in terms of language teaching and learning (Hughes & Hughes, 2020).

In the GSAT and TVE-JCEE English test, multiple-choice (MC) cloze tests are among the most common types of reading comprehension tests (Brown, 2002). This type of MC cloze test has been observed to be uncomplicated (Jonz, 1976). In MC cloze tests, students' English comprehension is tested by requiring them to select the best answer from four possible options to fill in the blanks in the passage to make a sentence semantically coherent and syntactically complete (Hao, 2011; Tabatabaei & Shakerin, 2013). Because several fundamental competencies are usually embedded in MC cloze tests, students are expected to fail to provide correct answers if their comprehension ability and logical thinking ability are not well developed (Luo, 2022). That is, several parts in different passages should be logical and comprehension clues that contribute to the meaningfulness of the whole passage; most students have found that this type of test is the most difficult of all exams. The use of contextual clues can be problematic for students, and a series of other difficulties may possibly co-occur (Katalin, 2000; Luo, 2022). Therefore, the need to investigate the issues and factors involved in performing and constructing MC cloze tests have begun to attract many scholars' attention (Bachman & Palmer, 2010; Chou & Chen, 2009; Tabatabaei & Shakerin, 2013).

Many studies have been done on these topics with different types of participants, including the exploration of the use of both senior and vocational high school students' (VHSs') strategies (Ai, 2015; Chen, 2013; Cheng, 2008; Joe, 1993; Kuo, 2003), the effects of scaffolding instructions on both senior and VHSs (Luo, 2022; Wang, 2018a), factors that affect junior high school and college students' performance on cloze tests (Azimi, 2016; Kumazawa, 2016; Trace et al., 2017), and features of the cloze test (Wang, 2018a). In addition, a few researchers (Kuo, 2003; Wang, 2018a) have found from observations that high school students encounter difficulties when taking MC cloze tests, although these conclusions were not based on scientific methods. Only a few studies have examined the difficulties that learners face from their own perspective. Although the reasons for both senior and VHSs' difficulties in performing well on cloze tests remain unexplored, the issue is more urgent for VHSs. Most VHSs are directed to gain specific skills due to the curriculum design, which is aimed at providing certification. Thus, confidence in learning English is gradually lost and huge discrepancies will appear in English competence as their peers at senior high schools continue to advance (Chang et al., 2007).

In Taiwan, there is a large amount of VHSs, after decades of this type of instruction (Chou, 1995; Xu, 1999). Currently 345,225 VHSs are studying in Taiwan, according to MOE statistics. In 2022, there were 79,292 VHSs taking the TVE-JCEE, only slightly less than the number of senior high school students taking the GSAT. Existing studies of test performance in vocational high school educational systems remain under-examined. Accordingly, this study investigated VHSs' difficulties in performing an MC cloze test, as well as whether any differences among the difficulties identified by VHSs. Finally, the factors that affect VHSs' performance difficulties in MC cloze test were examined and contrasted with the findings of previous studies. In particular, this study investigated the following research questions:

1. What types of difficulties do VHSs perceive in taking MC cloze tests?
2. What are the differences among the types of difficulties that are perceived by VHSs in taking MC cloze tests?

3. What factors affect VHSs' difficulties in taking MC cloze tests?

Through this study, it is hoped that significant and perspicacious implications will be derived in both the theoretical and pedagogical aspects. Theoretically, the difficulties that VHSs have in taking MC cloze tests should be given closer attention to studies of language assessment. From a pedagogical point of view, educators and test designers may understand what focuses should be brought to bear to promote students' testing strategies and performance in cloze tests. From this, better teaching procedures and curricula can be developed and designed. Most importantly, work along these lines will produce positive backwash, for the effects of tests on teaching and learning (Hughes & Hughes, 2020).

Literature review

Background of cloze test development

The cloze procedure, a technique used to assess text readability and communication effectiveness, was introduced by Wilson Taylor in 1953 (Bickley et al., 1970; Kumazawa, 2016). Unlike the testing concept of closure (Rankin, 1959), in which a missing gap is filled to complete a whole, as in Parviz and Sorayya (2012), the cloze procedure involves the systematic deletion of preselected texts to evaluate readers' competence by having them provide the precise words that were removed. From that point on, increasing interest in and attention to research on cloze procedure has been seen, including studies of the effectiveness of cloze test (Ajideh & Mozaffarzadeh, 2012; Akmedovna, 2022; Alderson, 1990), factors in cloze test performance (Tabatabaei & Shakerin, 2013), and item difficulty (Brown, 1989). Separate from these research topics, the use of the cloze procedure has become distinctive as a tool for conduct reliability research (Taylor, 1953). Results of such tools were considered to be diverse, examples of which were found in the reliability values, which ranged from 0.13 to 0.96 (Bachman, 1985; Brown, 1989; Pike, 1973), and criterion-related validity values, which ranged from 0.06 to 0.91 (Bachman, 1985; Brown, 1989). At the same time, a group of researchers began to focus their attention to the various types of cloze procedure, such as the C-test, developed by Raatz and Klein-Braley (1981), and MC cloze tests, developed by Jonz (1976). In addition to the two major types, several types of cloze procedure appeared, including a fixed-ratio cloze test (Cohen, 1994), a rational cloze test (Alderson, 2000), a conversational cloze (Brown, 1983), and a matching cloze (Baldauf & Propst, 1979). In addition, various scholars have held different points of view with respect to the types of cloze test. For example, Alderson (2000) considered the rational cloze to be a gap-filling test, while the random cloze type was restricted by the term cloze, meaning that it was only a low method to measure English proficiency. In addition, Bachman (1990) indicated that the types of cloze procedure should include rational deletion. Among the various classifications of cloze types, the MC cloze test is the only type that VHSs face in TVE-JCEE, so this study focuses on that.

Construction of the MC cloze test

Drawing on Goodman's (1967) psychological perspective, Boonsathorn (1987) developed the MC cloze test; the principle of the MC cloze test related to the belief in readers'

engagement of whole processing levels all at once. Due to the disadvantages of the C-test, it was expected that the MC cloze test could better test students' overall ability (Wonghirsombat, 2013). Regarding the MC cloze test, three critical aspects should be taken into considerations, including test passages, word deletion, and the distribution of testing points. Each aspect is presented and discussed in the following.

Text passages

First, text passages are crucial for constructing MC cloze tests (Ajideh & Mozaffarzadeh, 2012; Tavakoli et al., 2011). Let us take Tabatabaei and Shakerin (2013) as an example. The effectiveness of content familiarity on the cloze test performances of 60 Iranian EFL learners was investigated. A statistically significant difference was discovered between the testing results of MC cloze tests with familiar and unfamiliar content, where familiar content was linked to successful performance on the MC cloze tests. Likewise, Tavakoli et al. (2011) examined the effects of genre familiarity on an MC cloze test and a C-test. The results showed a significant impact of genre familiarity on both the MC cloze test and the C-test. In recent years, Trace (2023) investigated how the passage cohesion affected the function of the items. The results showed that the passage factors and item function are closely linked. The conclusion was made that aside from content and grammatical structure, test designers should investigate the impact of cohesion in potential cloze passages. Hughes and Hughes (2020) provided suggestions and measures to develop a relevant MC cloze test. First, the difficulty levels of selected passages should match the test takers' level of proficiency. After the issue of level is perfectly controlled, several passages should be involved in the trailing. Second, the text style should match with the level of language ability that is being tested. Third, as words are systematically deleted, it is critical to have native speakers closely inspect the test and provide their opinions on the ideally predetermined answers. Fourth, responses should be given with the provision of clear instructions, so that irrelevant factors can be diminished. Fifth, descriptions could be given to better interpret the scores on the MC cloze test. In the light of it, the text passage is an important factor in the perspective of constructing MC cloze test.

Deletions of words

For deletions of words in an MC cloze test, Cohen (1980) remarked that "A cloze test in its form is a passage from which after every certain number of words a word is deleted" (p. 91). However, Bachman (1985) believed that systematic and unsystematic deletions are both possible methods to use in making MC cloze test. From reviews of the existing literature, a systematic approach to deletion appears to be used more widely. In general, deletions are made on every n th word (Brown, 2002; Dhyaldian et al., 2022; Tabatabaei & Shakerin, 2013), and various words counts have been advocated the purpose of systematic deletions, including the deletion of roughly every fifth word (Yaseen & Rasheed, 2022), the deletion of every seventh word (Tavakoli et al., 2011), deletions of every sixth to eighth word (Tabatabaei & Shakerin, 2013), deletions between the fifth to tenth word (Azimi, 2016), deletions of every eighth or tenth word (Hughes & Hughes, 2020), and deletions of every twelfth word (Brown, 1989). Hughes and Hughes (2020) reported that in deletion, a few sentences at the beginning and in the end of the passages should

remain untouched so that any clues in this text can be referenced as test takers seek to complete the MC cloze test. In summary, MC cloze tests in TVE-JCEE appear to adopt the approach of deleting words based on a certain range, around 7 to 8 words on average between blanks. This measure is more reasonable because repeated or irrelevant testing constructs may still be included by applying sufficient exact word methods.

Distribution of testing points

The other critical aspect in constructing an MC cloze test is the development of relevant item constructs and sub-skills for testing. Due to constant changes in pedagogical beliefs, questions of what skills should be included and how far each construct should be incorporated have dynamically altered in terms of the accepted means of formulating MC cloze tests. Constructs can be equally divided into five items for the grammar aspect and the vocabulary aspect. Imbalanced testing constructs have been observed between the two aspects according to reviews of MC cloze tests on previous TVE-JCEE tests. Lu (2003) examined the item distribution across five consecutive years' items in MC cloze tests, ranging from 1998 to 2002. The questions mainly assessed test takers' comprehension within a relatively limited set of texts. The beliefs and conventions applied in making high-stake tests began to change as the 108 curricula were brought forward and preferred. After the implementation of these curricula, which stressed an orientation toward competency, the tendency to construct test items more comprehensively was perceived. From this educational policy, texts become longer, and test takers may need to apply more than one skill to complete the tests. Whether test takers' competencies are flawlessly assessed simply by increasing texts length and task complexity remains debatable and dubious, as test takers appear to be becoming more incompetent in completing the tests. In addition, concrete descriptions of students' test-taking difficulties have been blurred by the recent emergence of new test types.

Test takers' difficulty in performing cloze test

Many studies have been performed hitherto on MC cloze test difficulty, indirectly and obliquely indicating test takers' difficulties in performing this test (Abraham & Chapelle, 1992). Hughes and Hughes (2020) called for all tests to be carefully designed to allow test takers to know what to refer to. This indicates that an MC cloze test would be difficult with fewer clues. Boonsathorn (1987) determined the reliability of the C-test and the MC-Test using comparisons. Whether the different starting points of deletion would affect the difficulty was further explored. To investigate this, two forms of test, a C-test and an MC-Test, were created. Four tests were given to L1 and L2 participants. Both types of test were highly reliable. The MC-Test appeared to be more challenging for both L1 and L2 participants because the type of test required a greater than usual reading comprehension process, as well as better discrimination. In the same vein, Kumazawa (2016) inspected factors influencing score variance in MC cloze tests. In particular, the study investigated the linguistic and textual effects on an MC cloze test. The results identified interactions of those factors were found, and the reliability of MC cloze test was established. Although the primary goals of those two studies did not focus directly on test takers' difficulties in taking MC cloze tests, relevant ideas were implicitly revealed and inferred through the results.

For more direct results, Han (2022) investigated the relationships among vocabulary ability, use of vocabulary learning strategy, and cloze test performance. The participants were Korean college students. The results indicated a positive correlation between students' vocabulary ability and their performance on a cloze test. Although this study highlighted the importance of vocabulary competence, additional factors were uncovered, and the application of quantitative method prevented deeper insight. Most importantly, VHSs were ignored. In a study targeting participants VHSs, Wang (2018a, 2018b) reported that VHSs indeed had difficulty performing an MC cloze test. From her observations and teaching experience, the vocabulary of most VHSs was excessively limited, and they were unfamiliar with grammatical concepts and sentence patterns. For this reason, they were unable to comprehend the reading passages used in the MC cloze test. Most VHSs relied heavily on their rote learning, and they reported that there were too many targeted words and rules to remember. Ultimately, most VHSs described by Wang (2018b) gave up on learning English. Their frustrations and difficulties were vividly portrayed; however, it cannot be denied that evidence from observations may not match VHSs' inner thoughts. As the idea of a learner-centered approach to language assessment literacy has attracted significant attention recently (Butler et al., 2021; Lee & Butler, 2020), explorations of VHSs' test-taking difficulties can be scientifically conducted through the direct involvement of students as participants with proper instruments to elicit their inner voices.

Using learners' perspectives as a lens to investigate the difficulties of an MC cloze test, Ajideh and Mozaffarzadeh (2012) investigated whether the MC cloze test and C-test were appropriate to assess learners' reading comprehension. In addition, opinions and reflections on these two types of tests were further explored via a retrospective study. The results indicated that the MC cloze test was much more applicable for measuring test takers' reading comprehension than the C-test. For the results of participants' views of these two types of tests, it was found that the MC cloze test was easier to complete than the C-test, and these results are reasonable and predictable. Surprisingly, participants remarked that probability of guessing the correct answers was greater than 50% for both tests. Even if scientific methods were used to justify the results, involving advanced learners inevitably makes the results less convincing. Most importantly, it is urgent to explore VHSs' test-taking difficulties to provide timely support to them, as the results of TVE-JCEE English scores can be a decisive factor in being able to enroll in better colleges. Thus, the significance of exploring VHSs' difficulties in performing MC cloze test through learners' perspectives should be noted.

Method

Participants

The participants were Taiwanese VHSs studying at different grade levels. A total number of 309 VHSs completed the online questionnaire. Incomplete questionnaires and those where the same value was chosen for all items were discarded. After the deletion of 16 invalid or incomplete responses, 293 questionnaires were included in this study. In final group, there were 119 male students and 174 female students. In all, 73 students were in grade 10, 131 students were in grade 11, and 89 students were in grade 12. Most were studying at public vocational schools ($n = 250$), and the remainder students were at

Table 1 Demographic information of participants

Variables		Frequency	Percent
Gender	Male	119	40.6%
	Female	174	59.4%
Grade	Grade 10	73	24.9%
	Grade 11	131	44.7%
	Grade 12	89	30.4%
School type	Public	250	85.3%
	Private	43	14.7%
Total		293	100%

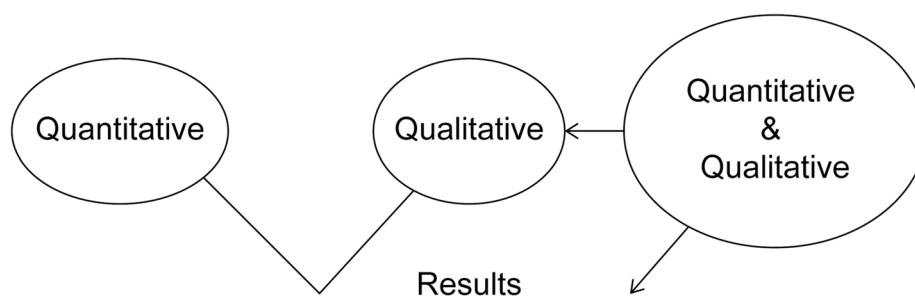


Fig. 1 Convergence model (Creswell, 1999)

private vocational schools ($n = 43$). Table 1 presents participants’ demographic information. These students mainly used textbooks published by San Min Book Co. or Longteng Education for English. All students had five English lessons a week, with each one lasting 50 min.

Research design

This study was a mixed-methods research. Creswell (1999) reported that convergent designs are triangulation designs, which feature both qualitative and quantitative data, and the results were generated by comparing the different types of data. In this manner, the results are powerful, taking account of the fact that the interpretation of data was integrated to justify the results (Caracelli & Greene, 1993). Accordingly, this study applied a convergence model to identify VHSs’ difficulties in performing well on MC cloze tests so that empirical reality is hoped to be comprehensively revealed (Creswell, 1999). Thus, a questionnaire consisting of 4-point Likert scales and a written narrative inquiry task was developed to achieve the purposes of this study. Figure 1 illustrates the convergence model.

Data collection method

A self-developed questionnaire was used to collect data to explore VHSs’ difficulties in performing an MC cloze test, as shown in the appendix. It is critical to have a well-designed study to obtain accurate results. Drawing on Krosnick and Presser (2010), this study’s questionnaire was created in three parts. The first part included 4-point Likert scales to explore VHSs’ MC cloze test difficulties. The second part was a written narrative

inquiry task to elicit VHSs' inner voices. The participants' demographic information was collected at the end of the questionnaire. For specific details of the construction of each part, the procedures and measures are described in the following paragraphs.

The questionnaire

The first part of the questionnaire collected data in the form of scoring on 4-point Likert scales. Before the construction of the first part, quick written interviews were carried out with two intact classes in central Taiwan. In all, 64 VHSs reported challenges met with in taking MC cloze tests. From the data collected in the interviews, seven facets were constructed based on common themes, including vocabulary, grammar, sentence structure, text length, text topic, clues designs, and test designs. These are presented in Table 2. Using 4-point Likert scales, it was hoped to eliminate the tendency to choose the middle value (Garland, 1991).

Written narrative inquiry

The second part of the questionnaire was a guided written task eliciting VHSs' inner thoughts. Instead of allowing free writing, three guided questions were created to enable the respondents to formulate their answers in a well-organized way. The guided questions required VHSs to identify the most difficult part concerning MC cloze test completions, their feelings with respect to those difficulties, and factors that may lead to these negative results.

Demographic information

The final part of the questionnaire gathered the participating VHSs' background information to better ground the results of the other two parts of the questionnaire. VHSs' sex, grade, and school type were collected. In addition, they were also asked whether they had chosen to study English. Some VHSs' responses were excluded because they had received particular training to study English, which could have altered the way that they viewed MC cloze tests could. This group take an English reading to evaluate their competences in their professional subject, and for this group, the MC cloze tests taken are more difficult than those taken in the general TVE-JCEE. Hence, VHSs studying in the English divisions were excluded from this study.

Table 2 Distribution of questionnaire items

Facets	Number of items
Vocabulary	3
Grammar	4
Sentence structure	4
Text length	2
Text topic	2
Clue designs	3
Test designs	6
Total	24

Data collection procedure

A complete data collection procedure was established using a convergence model. In general, there were five steps to the collection. First, quick written interviews were conducted on November 24, 2022, to generate ideas regarding VHSs' difficulty in MC cloze tests. Second, descriptions of items were discussed with an in-service teacher teaching at a public vocational high school in central Taiwan. Third, the questionnaire was uploaded online on December 4, 2022, after some minor modifications were suggested by an in-service teacher from Taichung. As this was an online questionnaire, descriptions of participants' rights were provided, and the participants volunteered to complete it online. Therefore, ethical soundness was not violated. Fourth, the online questionnaire was available for a month, and the link to the online questionnaire was made non-operational by the researcher on January 5, 2022. In the final step, invalid data were discarded, and the remaining data was transformed from Excel 2010 to SPSS 26 for further analyzing.

Data analysis

Both quantitative and qualitative data were collected, due to the convergent design of the present study. For the quantitative data, statistical packages in SPSS 26 were used to address research questions one and two. To answer research question 1, frequencies, means, and standard deviations were calculated and presented via descriptive statistics. One-way ANOVA was computed to determine whether there were interactions among the seven facets of VHSs' difficulties. Scheffé's test was utilized to explore details in interactions if statistically significant results were found. Scheffé's test is a powerful tool and sensitive to complex comparisons (Brown, 2005). With its use, research question 2 was sufficiently resolved. For VHSs' responses in the written narrative inquiry, thematic analysis was applied. Responses were grouped based on common themes, and interpretations of the students' texts were made to address research question 3.

Validity and reliability

A test–retest measure was conducted to validate the questionnaire. In all, 31 VHSs from eastern Taiwan participated in this process. Before the questionnaire was administered, informal consent was obtained from the students. The questionnaire was first administered on December 14, 2022, and then again on December 27, 2022. Because the questionnaire was designed to collect both quantitative and qualitative data, different methods were used to identify reliability were performed. For the quantitative data, SPSS 26 was used for reliability testing, and the results, shown in Table 3, were moderately high ($r=0.906$). For the qualitative data, an in-service English teacher who possessed a master degree in the field of TESOL was invited to thoroughly examine the collected data. It was found that the written narrative task effectively elicited VHSs' inner thoughts. The same teacher was invited to the main study to enable inter-rater reliability.

Table 3 The result of test–retest reliability

Cronbach's alpha	Cronbach's alpha based on standardized items	N of items
.906	.93	48

Results

This study was undertaken to identify the difficulties that VHSs encounter in taking MC cloze tests. To accomplish this goal, three research questions were addressed and used to systematically present the results and interpretations. Those research questions presented VHSs MC cloze difficulties, differences in perceived difficulties, and factors that affect the perceived difficulties. All of these are described and presented separately in the following.

VHSs' perceived difficulties in performing MC cloze test

As seen in Table 4, every facet's mean score was moderately high. Grammar had the highest mean score among the seven facets. Clue designs exhibited the second-highest mean score. Sentence structure had the lowest mean.

The frequencies of questionnaire items that were embedded in each facet were explored in detail. Percentiles for preferences were calculated by further computing the frequency numbers. Table 5 presents the frequencies of items in vocabulary. Most VHSs agreed that vocabulary used in the options, in the passages, and in collocations were difficult.

Grammar difficulties were investigated, and the results are presented in Table 6. Most VHSs chose strongly agree for all items, indicating that testing items related to tense, aspect, and part of speech were challenging. In addition, testing items related to judgments of voice and subjunctive mood were difficult.

Table 4 Descriptive information for VHS difficulties

Facets	N	Mean	Standard deviation
Vocabulary	293	2.88	0.83
Grammar		3.06	0.72
Sentence structure		2.64	0.86
Text length		2.83	0.68
Text topic		2.87	0.84
Clue designs		2.94	0.79
Test designs		2.81	0.78

Table 5 Response frequencies of items for vocabulary difficulty

Items	Strongly disagree	Disagree	Agree	Strongly agree
Q1	26 (8.9%)	82 (28%)	106 (36.2%)	79 (27%)
Q2	22 (7.5%)	61 (20.8%)	111 (37.9%)	99 (33.8%)
Q3	22 (7.5%)	73 (24.9%)	119 (40.6%)	79 (27%)

Table 6 Response frequencies of items for grammar difficulty

Items	Strongly disagree	Disagree	Agree	Strongly agree
Q1	11 (3.8%)	46 (15.7%)	85 (29%)	151 (51.5%)
Q2	20 (6.8%)	55 (18.8%)	100 (34.1%)	118 (40.3%)
Q3	24 (8.2%)	69 (23.5%)	99 (33.8%)	101 (34.5%)
Q4	27 (9.2%)	69 (23.5%)	89 (30.4%)	108 (36.9%)

Table 7 Response frequencies of items for sentence structure difficulty

Items	Strongly disagree	Disagree	Agree	Strongly agree
Q1	57 (19.5%)	117 (39.9%)	71 (24.2%)	48 (16.4%)
Q2	40 (13.7%)	106 (36.2%)	87 (29.7%)	60 (20.5%)
Q3	32 (10.9%)	95 (32.4%)	95 (32.4%)	71 (24.2%)
Q4	30 (10.2%)	56 (19.1%)	113 (38.6%)	94 (32.1%)

Table 8 Response frequencies of items for text length difficulty

Items	Strongly disagree	Disagree	Agree	Strongly agree
Q1	39 (13.3%)	126 (43%)	84 (28.7%)	44 (15%)
Q2	16 (5.5%)	41 (14%)	98 (33.4%)	138 (47.1%)

Table 9 Response frequencies of items for text topic difficulty

Items	Strongly disagree	Disagree	Agree	Strongly agree
Q1	24 (8.2%)	69 (23.5%)	123 (42%)	77 (26.3%)
Q2	23 (7.8%)	91 (31.1%)	109 (37.2%)	70 (23.9%)

Table 10 Response frequencies of items for clue design difficulty

Items	Strongly disagree	Disagree	Agree	Strongly agree
Q1	18 (6.1%)	31 (10.6%)	147 (50.2%)	97 (33.1%)
Q2	18 (6.1%)	75 (25.6%)	116 (39.6%)	84 (28.7%)
Q3	24 (8.2%)	86 (29.4%)	99 (33.8%)	84 (28.7%)

Table 7 presents the results of the frequency distribution in the facet of sentence structure. Most VHSs did not consider simple and compound sentences problematic as they were tested in the MC cloze test. Half of the VHSs disagreed, and half agreed that complex sentences were difficult. Compound-complex sentences were considered somewhat difficult, with the highest agreement rate in this section.

The results for the facet of text length difficulty are presented in Table 8. Disagree was selected by most VHSs, indicating that short text lengths tended to be easy to complete. However, most VHSs agreed that long text lengths were difficult for them to respond to.

Topic difficulties for texts are displayed in Table 9. Most VHSs chose agreement for both items. That is, when text topics were not related to the textbooks they had used or to topics they had in their daily lives, the MC cloze tests were challenging.

For the clue designs, the results of frequencies are presented in Table 10. Highly similar options led to difficulties. The clues were also difficult to locate. In addition, some VHSs did not understand the testing points. In general, agreement was selected by most VHSs.

Table 11 presents the results for response frequency. Most VHSs agreed that the blanks were too many and dense, so the MC cloze test was perceived to be difficult. The designs for the blanks were not easy to locate. In addition, most VHSs reported that the

Table 11 Response frequencies for items for test design difficulty

Items	Strongly disagree	Disagree	Agree	Strongly agree
Q1	27 (9.2%)	87 (29.7%)	108 (36.9%)	71 (24.2%)
Q2	28 (9.6%)	71 (24.2%)	122 (41.6%)	72 (24.6%)
Q3	29 (9.9%)	91 (31.1%)	103 (35.2%)	70 (23.9%)
Q4	20 (6.8%)	71 (24.2%)	118 (40.3%)	84 (28.7%)
Q5	30 (10.2%)	90 (30.7%)	98 (33.4%)	75 (25.6%)
Q6	23 (7.8%)	69 (23.5%)	113 (38.6%)	88 (30%)

MC cloze test mostly examined test takers’ rote memorization instead of comprehension; therefore, there were many vocabulary and grammar testing items in MC cloze tests.

Differences in perceived difficulties

One-way ANOVA was computed to examine differences in VHSs’ perceived difficulties among the seven facets. As shown in Table 12, significant differences were identified among perceived difficulties in performing the MC cloze test ($F(6, 2044) = 7.781, p < 0.001$).

Scheffé’s post hoc analysis was used to identify the details of the differences among the seven facets. In Table 13, a significant difference was found between the comparisons of grammar and sentence structure. Using sentence structure to compare vocabulary and clues designs, the results showed moderately significant differences. When sentence structure was compared to the text topic, a slightly significant difference was discovered, as was also the case for comparisons of test design and grammar.

Table 12 Results of one-way ANOVA

Source of variance	df	Sum of squares	Mean square	F
Between-groups	6	29.32	4.88	7.781***
Within-group	2044	1283.79	0.628	
Total	2050	1313.12		

Note: *** $p < .001$

Table 13 Results for Scheffé’s post hoc analysis

(I) Group	(J) Group	Mean difference (I-J)	Std. error	Sig
Vocabulary	Sentence structure	.244	.065	.004
Grammar	Sentence structure	.421		.000
	Test design	.253		.021
Sentence structure	Vocabulary	-.244		.03
	Grammar	-.421		.000
	Text topics	-.234		.046
	Clue design	-.304		.002
Text topics	Sentence structure	.234		.046
Clue design	Sentence structure	.304		.002
Test design	Grammar	-.253		.021

Comparisons for which no significance differences were found are not presented. Two primary orders of VHSs' perceived difficulties in completing MC cloze tests can be generated:

1. Sentence structure was reported to be the least difficult facet.
2. Grammar was more difficult than test design.

In addition to the quantitative data, a written narrative inquiry was used to collect VHSs' qualitative data. Three questions were developed to serve as guidance for VHSs to respond in a well-organized way. The first question called on VHSs to choose the most difficult facet among the seven categories. After classification and discussion with an in-service teacher in the field of TESOL, most responses were seen to be related to the recognition and comprehension of vocabulary and grammar. Some VHSs noted that collocations related to the uses of prepositions were difficult. Some examples of this are presented illustrated in Table 14:

To support VHSs' responses for the first written narrative inquiry question and further clarify the perceived difficulties that VHSs had for the MC cloze test, the second question of the written narrative inquiry task required VHSs to express their feelings toward such difficulties. In general, negative feelings were reported. In particular, three main themes were identified in the analysis of the VHSs' responses. First, several VHSs indicated that they had given up on learning because MC cloze tests are too hard. They did not want to read the texts, and they did not want to answer the questions. Second, some of the VHSs found it frustrating and felt helpless when faced with such difficulties. In addition, they felt anxious and annoyed when working on MC cloze tests. Third, some students responded that they saved the MC cloze questions to the last minute during their tests. That is, they considered that they could make better use of their time in responding to questions that they were certain they could correctly answer, such as reading comprehension. Some examples of this are given in Table 15.

Table 14 The most difficult facet for VHSs in doing MC cloze test

S 25: (Translation)	跨某單字。 The vocabulary is incomprehensible
S 46: (Translation)	不知道時態。 The tenses are unknown
S 98: (Translation):	文法有些較難。 Some grammar concepts are more difficult
S 157: (Translation):	單字太深奧。 Vocabulary is too recondite
S 174: (Translation)	單字都看不懂。 Not all of the vocabulary is understandable
S 261: (Translation)	單字量不夠, 分不出時態。 My vocabulary size is limited. Tenses and aspects are hard to distinguish
S 276: (Translation)	記不住單字。 The vocabulary is hard to remember

Table 15 VHSs' feelings toward difficulties in the MC cloze tests

S 3:	很焦慮。
(Translation)	Very anxious
S 10:	會不想學... 因為都學不會。
(Translation)	I don't feel like learning... because the skills cannot be acquired
S 34:	不會想讀。
(Translation):	I don't want to study English
S 54:	我放棄。
(Translation):	I give up
S 61:	很挫敗。
(Translation)	It was frustrating
S 72:	會慌亂 因為有很多不懂的。
(Translation)	I will feel panic because there are many things I don't know
S 95	好煩 留到最後在寫。
(Translation)	It is annoyed. I leave it to the end and respond then
S 116:	越學越累。
(Translation)	The more I learn, more fatigued I feel
S 138:	看到就想放棄。
(Translation)	As soon as I see them, I want to give up
S 262:	習慣了, 我只能把克漏字留到最後, 有把握的拿分。
(Translation)	I am used to it. I save it for the end and answer some other parts that I am quite certain of
S 286:	先跳過寫會的最後在重頭來看。
(Translation)	I will skip it and answer other questions that I know. In the end, I will go back to it to read from the beginning
S 292:	很懊惱。
(Translation)	It is upsetting

VHSs' self-perceiving factors in MC cloze test difficulties

In the written narrative inquiry task, the third question asked VHSs to self-report factors that had caused difficulty in completing MC cloze tests. Five main themes were generated. First, most VHSs reported that they had vocabularies that were too small. Because they tended to forget words that they had memorized, they found it hard to comprehend reading passages and options. In addition, they reported wishing to read more authentic materials written in English so that they would have the opportunity to learn words. Second, some VHSs revealed that they wished to put more effort into learning and reviewing English grammar. As they believed that they did not lay a good foundation for themselves while learning, it was hard for them to choose the correct answers. Third, some VHSs blamed themselves for poor intelligence or poor memory. Fourth, some VHSs found faults with their teachers' teaching style and the difficulty of MC cloze-style testing. Finally, a few VHSs reported that they were not consciously aware of any factors that affected difficulties they might have. Representative VHSs are given in Table 16.

Discussions

The quantitative and qualitative analyses generated three critical points, as follows. First, most VHSs reported difficulty in learning vocabulary and grammar. They wished that they could put more efforts in learning those two basic components.

Table 16 Perceived factors of MC cloze test difficulties

S 18:	單字量不多。
(Translation):	I do not have enough vocabulary
S 26:	老師故意出很難。
(Translation):	Teachers intentionally make them difficult
S 40:	老師的問題。
(Translation):	It is teachers that cause these problems
S 54:	英文一個意思可以用太多單字了背不起來。
(Translation):	There are many synonyms so that I find them hard to remember
S 59:	本身文法要加強。
(Translation):	I have to strengthen my English grammar
S 81:	單字量不夠, 文法不熟練。
(Translation):	I have too little vocabulary, and I am unfamiliar with the grammatical concepts
S 115:	記憶力不好。
(Translation):	My memory is poor
S 168:	我太笨。
(Translation):	I am too stupid
S 173:	頭腦有問題。
(Translation):	There are problems with my brain
S 247:	老師不會教, 只知道叫我們背, 老師愛酸學生。
(Translation):	The teachers do not know how to teach, and they demand that they memorize everything. They like to tease students
S 263:	不知道。
(Translation):	I do not know
S 280:	自己英文讀的不夠多。
(Translation):	I do not read many things about English

This contradicts Hughes and Hughes's (2020) concepts of clue design in determining VHSs' MC cloze test difficulties. In addition, VHSs reported difficulties appear to contradict trends in English teaching in Taiwan, which seeks to be more comprehensive and applicable. These results line up with Kumazawa (2016) finding that indicated vocabulary plays a decisive role in MC cloze test difficulties. In the same vein, the results are consistent with Han (2022), suggesting a positive correlation between vocabulary competence and cloze test performance. The results of this study does not support Yaseen and Rasheed (2022), who indicated that cloze testing is an appropriate means of assessing students' language competency. The reason for such outcomes is clear, as VHSs cannot move on to the next level of learning without basic competence in the previous level. When VHSs attended vocational high schools instead of senior high schools, they are assumed to possess less than basic English competence. It is suggested to provide more assistance to solidly build their English competence as well as their confidence. Besides, after years test-taking experiences, VHSs do not seem to gain required English proficiency, showing that the concept of assessment for learning is not fulfilled (Hughes & Hughes, 2020). In this regard, mountainous and monotonous tests may have cause negative effects on VHSs learning processes. Appropriate and diverse modifications of test format are needed to better meet the needs of VHSs (Jin, 2023).

Second, sentence patterns were the least difficult part as perceived by VHSs when performing the MC cloze test. This indicates that there may be something wrong with the test designs since scholars alike believe in the notion that the sentence pattern should be embedded inside English grammar. With this scholar-centered or teacher-centered concept, the insufficient considerations of prioritizing learning sequence surface on the table. In addition, the VHSs expressed the belief that the testing items in the MC cloze test concerning English grammar are not clearly designed. In other words, the clues are not sufficient, and they do not know how to respond to those questions correctly, so long as their grammatical concepts are not solidly learned. This idea seems to violate the concept of assessment for learning (Hughes & Hughes, 2020). To resolve the problem, Weideman and Dyk (2023) investigated a format that is rarely seen and applied in Taiwan, shedding lights on possible directions of future cloze test modification in a more effective and interdependent fashion.

Third, negative thoughts are identified after the data analyses. It is highly possible that teachers' teaching styles and their roles greatly affect VHSs' already weak learning motivations since various factors contributing to learning outcomes are likely influenced by these teaching approaches (Hughes & Hughes, 2020). This issue demonstrates that there exists an emergency to promote deeper understanding of VHSs learning conditions. Although novel teaching pedagogy has constantly been the target of advocacy and experimentation, such as CLT, TBI, and CLIL, based on the VHSs' responses, some teachers still use traditional approaches that have led students to dislike English study. In the view of this, the unchanged cloze test format becomes a focal point. Teachers, accustomed to regular teaching routines, may resist changing their approaches to better suit students' needs. Recognizing the importance of aligning teaching methods with VHSs' needs, changing the test format emerges as a crucial and initial step toward promoting a more student-centered approach. With regard to Taiwan's inherently test-oriented educational system, the implications drawn from a comprehensive understanding of VHSs' learning conditions become crucial (Jin, 2023). These insights serve as the groundwork for potential modifications to cloze tests, strategically guiding teachers' instructions toward a desired and student-centric manner.

Conclusion and suggestions

This study investigated VHSs' difficulties in performing an MC cloze test, and factors that affect these difficulties were explored. Both quantitative and qualitative approaches were used to collect the data. While the findings only showed the results for VHSs' perceived difficulties and self-aware factors, two profound implications can be drawn, from a theoretical and pedagogical point of view. This study identified the significance of using learners' perspectives as a lens to examine students' difficulties in taking MC cloze tests. This allows mismatches between teachers' and learners' beliefs on assessment to be explicitly shown, and the idea of assessment for learning or positive backwash effects can be better fulfilled through identification of essential solutions. In addition, the research scopes of the relevant topics can be widened as learners tend to be versatile. Through an exploration of students' perspectives, theories test making can be revisited and improved instead of remaining scholars' wishful thinking only. Regarding some advice on changes of test designs, it is suggested to have more texts but fewer testing questions are suggested because VHSs indicated

that longer texts are more challenging. In addition, questions can be designed to require or elicit VHSs reasons for choosing the options to allow the goals of comprehensiveness. Means of testing vocabulary can be elegantly designed by asking VHSs for clarification or even paraphrasing. Even better, suggested by Weideman and Dyk (2023), the cloze test can be designed into asking VHSs to indicate the missing words and their locations. In terms of pedagogical implications, VHSs' learning motivations should be seriously taken into considerations. In addition, incorporating proper tasks to enable VHSs to develop their long-term memory areas can be critical to prevent them from forgetting what they have been taught. Most importantly, periodic energizing and evaluating teachers' growths of instructional abilities should be thoroughly organized and implemented.

Although this study offered insights into VHSs' perceived difficulties, there are two important limitations that can be used as directions for future study. On the one hand, the sample size of the present study was relatively small, and the majors of participants were not revealed. Thus, the results could not be generalized to the broader population. On the other hand, the interactions among difficulties were indicated and further explored, but more detailed and deeper investigation into learners' perspectives should be undertaken. Thus, more insightful suggestions can be established for educators' and test makers' reference. Moreover, the modifications of MC cloze test based on learners' opinions require extra validation in terms of reliability.

Abbreviations

VHS	Vocational high student
MC	Multiple choice
CLT	Communicative language teaching
TBI	Task-based instruction
CLIL	Content and language integrated instruction

Acknowledgements

The author wishes to express the gratitude to all the participants who kindly helped with data collection in this study. In addition, the author also wants to give thanks to the teacher who helped in the data analysis process.

Authors' contributions

The author completed the present study on his own, including concept development, writing up the research paper, data collection and analysis, and approved the final manuscript.

Funding

No funding was gained when the present study was completed.

Availability of data and materials

Both of the qualitative and quantitative data collected and analyzed in the present study can be obtained from https://drive.google.com/drive/folders/19-IFMS9c_6jtg5UQuqpc4K-QjsOIEfsy?usp=sharing.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

The author agrees the publication once the study is accepted.

Competing interests

The authors declare that they have no competing interests.

Received: 1 November 2023 Accepted: 21 January 2024

Published online: 26 April 2024

References

- Abraham, R. G., & Chapelle, C. A. (1992). The meaning of cloze test scores: An item difficulty perspective. *Modern Language Journal*, 76(4), 468–479. <https://doi.org/10.1111/j.1540-4781.1992.tb05394.x>
- Ai, W. C. (2015). Senior high school students' test-taking strategies on multiple-choice cloze tests and the relationship with English proficiency [Master's Thesis]. Southern Taiwan University of Science and Technology.
- Ajideh, P., & Mozaffarzadeh, S. (2012). C-test vs multiple-choice cloze test as tests of reading comprehension in Iranian EFL context: Learners' perspective. *English Language Teaching*, 5(11), 143–150. <https://doi.org/10.5539/elt.v5n11p143>
- Akmedovna, B. M. (2022). The effectiveness of cloze tests in assessing the reading of B2 level learners. *International Journal on Integrated Education*, 5(6), 329–336.
- Alderson, J. C. (1990). Testing reading comprehension skills (Part two): Getting students to talk about taking a reading test (A Pilot Study). *Reading in a Foreign Language*, 7(1), 465–503.
- Alderson, J. C. (2000). Assessing reading. Cambridge University Press. <https://doi.org/10.1017/CBO9780511732935>
- Azimi, M. (2016). The relationship between anxiety and test-taking C-test and cloze test. *Malaysian Online Journal of Educational Sciences*, 4(1), 30–42.
- Bachman, L. F. (1985). Performance on cloze tests with fixed-ratio and rational deletions. *TESOL Quarterly*, 19(3), 535–555. <https://doi.org/10.2307/3586277>
- Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford University Press.
- Bachman, L. F., & Palmer, A. S. (2010). *Language testing in practice*. Oxford University Press.
- Baldauf, R. B., & Propst, I. K. (1979). Matching and multiple-choice cloze tests. *Journal of Educational Research*, 72(6), 321–326. <https://doi.org/10.1080/00220671.1979.10885183>
- Bickley, A. C., Ellington, B. J., & Bickley, R. T. (1970). The cloze procedure: A conspectus. *Journal of Reading Behavior*, 2(3), 232–249. <https://doi.org/10.1080/10862967009546900>
- Boonsathorn, S. (1987). C-tests, proficiency, and reading strategies in ESL. [Doctoral Dissertation]. University of Alberta.
- Brown, J. D. (1983). A closer look at cloze: Validity and reliability. In J. W. Oller Jr. (Ed.), *Issues in language testing research* (pp. 237–250). Newbury House.
- Brown, J. D. (1989). Cloze item difficulty. *Journal of the Japan Association of Language Teachers*, 11(1), 46–67.
- Brown, J. D. (2002). Do cloze work? Or is it just an illusion? *Second Language Studies*, 21(1), 79–125.
- Brown, A. M. (2005). A new software for carrying out one-way ANOVA post hoc tests. *Computer Methods and Programs in Biomedicine*, 79(1), 89–95. <https://doi.org/10.1016/j.cmpb.2005.02.007>
- Butler, Y. G., Peng, X. L., & Lee, J. Y. (2021). Young learners' voices: Towards a learner-centered approach to understanding language assessment literacy. *Language Testing*, 38(3), 429–455. <https://doi.org/10.1177/0265532221992274>
- Caracelli, V. J., & Greene, J. C. (1993). Data analysis strategies for mixed-method evaluation designs. *Educational Evaluation and Policy Analysis*, 15(2), 195–207. <https://doi.org/10.3102/01623737015002195>
- Chang, W. C., H. N. Yeh, S. G. Joe, Y. L. You, C. L. Chern & M. L. Liaw (2007). Research on English-in-education policies and their effects on English teaching]. *The Proceedings of 2007 International Conference and Workshop on TEFL & Applied Linguistics*. Taipei: Crane, 672–686.
- Chen, T. T. (2013). The study and analysis of comprehensive school English Majors' test-taking strategies on multiple-choice cloze test [Master's Thesis]. National Taiwan University of Science and Technology.
- Cheng, H. Y. (2008). *The effects of test-taking strategy instruction on ELF learners' performance on cloze test: A case study low achievers in vocational high school* [Master's Thesis]. National Taiwan Normal University.
- Chou, Z. T. (1995). *An analysis of the criteria of test designing of the English test of Join High School Entrance Examination [Paper presentation]*. Crane Publishing Co Ltd.
- Chou, S. Y., & Chen, Y. M. (2009). A study of cloze test items in scholastic aptitude English test and department required test [Conference presentation]. 26th International Conference on English Teaching and Learning in the R. OC.
- Cohen, A. (1980). Testing language ability in the classroom. Newbury house.
- Cohen, A. D. (1994). Assessing language ability in the classroom. *Heinle ELT*.
- Creswell, J. W. (1999). Mixed-method research: Introduction and application. In G. Cizek (Ed.), *Handbook of education policy*. Academic.
- Dhyaaldian, S. M. A., Al-Zubaidi, S. H., Mutlak, D. A., Neamah, R., Albeer, A. A. M., Hamad, A. D. A., Hasani, S. F. A., Jaber, M. M., & Maabreh, H. G. (2022). Psychometric evaluation of cloze tests with the Rasch Model. *International Journal of Language Testing*, 12(2), 95–106.
- Garland, R. (1991). The midpoint on a rating scale: Is it desirable. *Marketing Bulletin*, 2(1), 66–70.
- Goodman, K. S. (1967). Reading in Psycholinguistic guessing game. In F. V. G. Goodman (Ed.), *Language and literacy, 1. The selected writings of Kenneth Goodman* (pp. 33–43). Routledge & KeganPual.
- Han, M. (2022). EFL learners' vocabulary ability, vocabulary learning strategies and performance of the cloze test. *Studies in Language*, 38(2), 143–156.
- Hao, L. (2011). Brief analysis of cloze test in high school. *Journal of Shaanxi Normal University*, 199–200.
- Hughes, A., & Hughes, J. (2020). *Testing for language teachers* (3rd ed). Cambridge University Press.
- Huttner, J. (2008). The genres of student writing: Developing writing models. *International Journal of Applied Linguistics*, 18(2), 147–165. <https://doi.org/10.1111/j.1473-4192.2008.00200.x>
- Jin, Y. (2023). Test-taker insights for language assessment policies and practices. *Language Testing*, 40(1), 193–203. <https://doi.org/10.1177/02655322221117136>
- Joe, S. G. (1993). A preliminary investigation of reading strategies on cloze tasks for high school students in Taiwan [Conference presentation]. The tenth conference on English teaching and learning, Taipei, Taiwan.
- Jonz, J. (1976). Improving on the basic egg: The M-C cloze. *Language Learning*, 26(2), 255–265. <https://doi.org/10.1111/j.1467-1770.1976.tb00276.x>
- Katalin, B. (2000). Reflections on the test-taking strategies of 7th and 11th grade Hun Garian students of English. *ELT Journal*, 7(3), 48–59.
- Kohonen, V. (1999). Authentic assessment in affective foreign language education. In J. Arnold (Ed.), *Affect in language learning*. Cambridge University Press.
- Krosnick, J. A., & Presser, S. (2010). *Handbook of survey research*. Emerald Group Publishing.

- Kumazawa, T. (2016). Factors affecting multiple-choice cloze test score variance: A perspective from generalizability theory. *International Journal of Language Studies*, 10(1), 15–30.
- Kuo, W. C. (2003). *Differences in processing tactics on cloze tests between successful and less successful readers: A case study* [Master's Thesis]. National Kaohsiung Normal University.
- 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. <https://doi.org/10.1191/0265532204lt260oa>
- Lee, J., & Butler, Y. G. (2020). Reconceptualizing language assessment literacy. Where are language learners? *TESOL Quarterly*, 54(4), 1098–1111. <https://doi.org/10.1002/tesq.576>
- Lu, H. Y. (2003). *A study of EEFTC English cloze procedure and reading comprehension test* [Mast Thesis]. National Yunlin University of Science and Technology.
- Luo, Y. (2022). The application of scaffold to college entrance examination—Cloze test from 2018–2020. *The Educational Review*, 6(6), 263–267. <https://doi.org/10.26855/er.2022.06.010>
- Madsen, H. S. (1983). *Techniques in testing*. N.Y.
- Parviz, A., & Sorayya, M. (2012). C-test vs. multiple-choice Cloze Test of reading comprehension in Iranian EFL context: Learners' perspective. *English Language Teaching*, 5(11), 1916–4742. <https://doi.org/10.5539/elt.v5n11p143>
- Pike, L. W. (1973). *An evaluation of present and alternative item formats for use in the Test of English as a Foreign Language*. Educational Testing Service.
- Raatz, U., & Klein-Braley, C. (1981). The C-test—A modification of the cloze procedure. In T. Culhane, C. Klein-Braley, & D. Stevenson (Eds.), *Practice and problems in language testing* (pp. 113–148). University of Essex.
- Rankin, E. F. (1959). *The definition of reading comprehension*. University of Minnesota.
- Tabatabaei, O., & Shakerin, S. (2013). The effect of content familiarity and gender on EFL learners' performance on MC cloze test and C-test. *International Journal of English Language Education*, 1(3), 151–171. <https://doi.org/10.5296/ijele.v1i3.3952>
- Tavakoli, M., Ahmadi, A., & Bahrani, M. (2011). Cloze test and C-test revisited: The effect of genre familiarity on second language reading test performance. *Iranian Journal of Applied Linguistics (IJAL)*, 14(2), 173–204.
- Taylor, W. (1953). Cloze procedure: A new tool for measuring readability. *Journalism Quarterly*, 30, 414–433.
- Trace, J. (2023). The influence of passage cohesion on cloze test item difficulty. *Language Teaching Research Quarterly*, 37, 161–178. <https://doi.org/10.32038/ltrq.2023.37.08>
- Trace, J., Brown, J. D., Janssen, G., & Kozhevnikova, L. (2017). Determining cloze item difficulty from item and passage characteristics across different learner backgrounds. *Language Testing*, 34(2), 151–174. <https://doi.org/10.1177/0265532215623581>
- Wang, J. (2018a). Features and strategies of cloze test in college entrance examination. *Education Science Forum*, 7, 47–50.
- Wang, Y. T. (2018b). *The effects of scaffolding strategy on vocational high school students' performance on cloze test* [Masters Thesis]. National Taiwan Normal University.
- Weideman, A., & Dyk, T. V. (2023). Achieving technical economy: A modification of cloze procedure. *Language Teaching Research Quarterly*, 37, 144–160. <https://doi.org/10.32038/ltrq.2023.37.07>
- Wonghiramsombat, P. (2013). The C-test and the MC-test. *Thammasat Review Special Issue*, 189–202.
- Xu, H. L. (1999). *The reflection and perspective of Entrance Examination for 4-year Technological Colleges* [Paper presentation]. The ninth International Symposium on English Teaching (pp. 627–638). Crane Publishing Co, Ltd.
- Yaseen, A. A., & Rasheed, W. A. (2022). Investigating EFL university students' written performance on cloze test. *Journal of Positive School Psychology*, 6(10), 1273–1281.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.