

What Is the Optimum Length of a Cloze Test?

Mohammadtaghi Shahnazari-Dorcheh¹, Saeed Roshan² & Akbar Hesabi¹

¹ Department of English Language and Literature, University of Isfahan, Isfahan, Iran

² University of Auckland, Auckland, New Zealand

Correspondence: Mohammadtaghi Shahnazari-Dorcheh, Department of English Language and Literature, University of Isfahan, Isfahan, Iran. E-mail: m_t_shah@yahoo.com

Received: August 23, 2012

Accepted: September 10, 2012

Online Published: September 26, 2012

doi:10.5539/ijel.v2n5p142

URL: <http://dx.doi.org/10.5539/ijel.v2n5p142>

Abstract

Following Nation' (2009) proposal of 40-50 empty spaces as an optimum length of a cloze test, this study examined whether this length would work according to proficiency level. Three cloze tests, adjusted for each proficiency level, were developed by the researcher for the use at beginning, intermediate and advanced levels. The cloze tests measured participants' reading comprehension and included 40, 45 and 50 empty spaces for the beginning, intermediate and advanced levels respectively. Problematic items of each cloze test were identified over several pilot studies. Then they were administered to three groups of L1 Persian EFL learners at the beginning (56), intermediate (43), and advanced (41) levels. Results of the study suggest that an optimum length of a cloze test could vary according to the proficiency level. While the test could be long (including 50 empty spaces) and reliable at the advanced level, a shorter length of the test including 20-25 empty spaces could be more reliable at the beginning and intermediate levels.

Keywords: cloze test, reading comprehension, proficiency, reliability, EFL learners

1. Introduction

Brown (1998) argues that at least three factors influence the reliability of cloze tests. These factors include variations in students' proficiency levels, differences in readability of the cloze tests, and length of the test (changes in number of items). Research suggests that if the first two factors are held constant, longer cloze tests yield more reliable estimates than shorter ones (e.g., Bachman, 1990; Brown, 1996, 1998). However, these researchers have not clearly determined the optimum length of a cloze test. Rand (1978) examined the effect of test length on the reliability of cloze tests. He suggested that a cloze test should include 20-25 empty spaces irrespective of the method used to score that. However, he had drawn on a single cloze test and single population in his study. Then variations in his participants' language proficiency and potential differences in readability of cloze tests were simply ignored in his study. Nation (2009) believes that one of the factors that make a cloze test reach a good level of reliability is that the cloze has a high number of points of measurement. He proposes that a cloze test with 40-50 empty spaces would be more reliable. However, it has not been argued whether the proposed length would work across different proficiency levels. Brown (1998) believes that there is a threshold at which increasing length of the test can have opposite effect. He has proposed a procedure through which one can determine the appropriate length of a cloze test. However, it would be very time-consuming for EFL/ESL teachers to determine the optimum length of a cloze test for each group of their participants. This is because the participants may differ in terms of language proficiency. To tackle this problem, it appears that a framework for an optimum length of a cloze test at each proficiency level is required. To meet this requirement, this study was designed to examine (a) whether the optimum length of a cloze test varies according to proficiency level, (b) if so, what would be the optimum length for a cloze test at each proficiency level.

1.1 Cloze Test

A cloze test is defined as a passage with a few sentences of introduction followed by text with deleted words (gaps) with a consistent number of words (from five to eleven) between them. The test taker's responsibility is to predict

the deleted words based on the words given in the text (e.g., Alderson, 2000; Brown, 1989, 1992; Cain & Oakhill, 2006; Koda, 2005; Nation, 2009; Oller, 1975, 1979). The rationale behind a cloze test is that the reader must be sensitive to both semantic and syntactic constraints in each context to fill in the blanks. Such sensitivity can be taken as a reliable indicator of reading ability since text information processing is supported by these constraints (Koda, 2005, p. 239).

Cloze tests can be constructed in two ways, by standard fixed ratio deletion or rational deletion format (Koda, 2005). In standard fixed ratio deletion format, every 5th or 7th word is deleted from the text. The missing word can be from any category of function or content words. However, the standard fixed ratio format has some drawbacks: such clozes are more sensitive to surface linguistic forms than the underlying text meaning, they say little about global-text information processing competencies, and may reflect memory performance rather than reading ability (e.g., Alderson, 2000; Beach, 1997; Cain & Oakhill, 2006; Koda, 2005). For these reasons, a rational deletion format was preferred. In this test, there is no fixed ratio to delete words from the text; rather, a prespecified category of words is deleted to test a specific facet of reading ability. Therefore, to measure text-meaning understanding, a proportion of content words are chosen and deleted. Contextual support availability rather than the number of deleted content words will determine the rational cloze test's difficulty level (Koda, 2005, pp. 240-241).

A multiple-choice format, rather than a fill in the blank one, was used in this study. This was to improve scoring reliability. Moreover, it was more practical to administer (e.g., Bachman, 1990; Heaton, 1988; Madsen, 1983) and a more familiar format for the participants. It consisted of 40 to 50 empty spaces where the readers' task was to read the passages and choose a replacement word from a choice of 4 options (Alderson, 2000; Cain & Oakhill, 2006; Koda, 2005).

2. Methodology

2.1 Subjects

A total number of 140 L1 Persian EFL learners, at beginning (56), intermediate (43) and advanced level (41) participated in the study. They included both males and females, 16-35 years old, studying English as a foreign language in a private language school.

Identification of the proficiency level of the participants was already done upon their enrolment. It was based on the New Interchange/Passages Placement test (Lesley, Hansen, & Zukowski/Faust, 2003), a test used in the language school where the participants were selected. Furthermore, a modified version of the Placement test was administered to participants of this study to ensure that each participant was placed at the right proficiency level. This PlacementTest was a 70-item multiple-choice test that lasted 50 minutes. However, to obtain an accurate proficiency score and not to allow the variable of time constraint to influence the results, an additional 40 minutes were given to the participants. Overall, 90 minutes were given for this test, although most of the participants completed it within the range of 60 to 80 minutes.

The modified version of Placement test measured L2 learners' listening, reading, and grammar recognition of English. With the exception of two participants, the results were consistent with the original proficiency placement. To score the test, one mark was allocated to the participants' correct answer (total of 70). The participants were placed at an appropriate proficiency level depending on their scores in this test. Scores of 30, 49, and 70 were determined as the cut-off scores for the beginning, intermediate and advanced levels respectively. A one-way ANOVA was also run to see if there were any significant differences between the assigned groups. The results of this analysis indicated that there were significant differences between the three proficiency groups, $F(2, 137) = 512.952, p = .000$. More specifically, the one-way ANOVA analysis indicated that each proficiency group differed from the other two groups significantly and with a large effect size ($\eta^2 = .88$). Table 1 indicates the descriptive statistics of this analysis for each proficiency level and how the participants were distributed into three levels of proficiency based on the Objective Placement Test scores.

Table 1. Descriptive statistics for the language proficiency scores for each proficiency level

	Beginning	Intermediate	Advanced
N	56	43	40
Gender	M= 21, F= 35	M= 14, F= 29	M= 5, F= 35
Age Range	16-27	16-31	18-35
Years of Study	.5-1	1.5-2	3-3.5
Min	6	31	50
Max	29	49	65
Mean	20.08	39.06	55.68
S.D	5.70	5.60	4.88

Note. M = Male; F = Female

2.2 Materials

A multiple-choice reading cloze test of 40-50 minutes in duration was used at each proficiency level. They measured participants' reading comprehension. The length of each cloze test differed depending on the proficiency level. They included 254, 318 and 370 words for the beginning, intermediate, and advanced levels respectively. To control the difficulty level of the reading passages at each proficiency level, Cobb's (2002) VocabProfiler, which classifies the words by word family, type, and token, was used. This software indicates how many words each passage contains from the following four frequency levels proposed by Laufer and Nation's (1995) lexical frequency profiler: (a) the list of the most frequent 1000 word families (K1), (b) the second 1000 (K2), (c) the academic word list, and (d) words that do not appear on the other lists (off-list). Table 2 displays vocabulary information for each cloze passage.

Table 2. Readability and vocabulary information of the cloze passages for three proficiency levels

Vocab Information	Beginning Level	Intermediate Level	Advanced Level
K1	83.10%	77.80%	72.62%
K2	2.61%	1.04%	4.38%
K1+K2	90.34%	85.01%	80.07%
AWL	0.39%	3.66%	3.92%
Off-List	1.74%	3.82%	13.04%
Readability	68.40	58.50	52.70

As indicated in Table 2, while the proportion of the academic word list (AWL) and off-list is higher for the advanced level, the proportion of the first and second thousand level words (K1+K2) is higher for the beginning level with the intermediate level in between. This suggests that the difficulty level of the passages increases for each proficiency level as the proportion of most frequent words decreases. Second, the density of the propositions in the passages was controlled. It is a variable which affects understanding and recall (Alderson, 2000; Nation, 2009; Beach, 1997; Koda, 2005; Cain & Oakhill, 2006). To do this, the total number of major and minor idea units in each sentence was determined. These were added up to return the total number of major and minor idea units per passage. This included 35, 43 and 49 idea units for the beginning, intermediate and advanced levels respectively.

A pilot study was conducted among the participants at each proficiency level to gain the participants' feedback on the appropriateness of a sample of the selected passages. Since the participants may have discussed similar content in class time, they were asked to read through each passage and fill out a retrospective report. In this report, the participants were asked to determine the extent to which they had already been familiar with the content of the passages, on a Likert scale of 4 items (very familiar, rather familiar, less familiar, and unfamiliar). Their reports

indicated that 100% of the participants had been unfamiliar with the content of the passages for the three proficiency groups.

2.2.1 Developing and Piloting the Cloze Tests

At each proficiency level, an appropriate passage (as described above) was chosen, and 40-50 content words were deleted. The number of deleted words differed for each proficiency level depending on the length of each text. There were 40, 45 and 50 deleted words in the selected passages for the beginning, intermediate and advanced levels respectively. The difficulty level was also controlled by leaving sufficient and equal contextual support for the learners to restore the deleted content words (Abraham & Chapelle, 1992). This was done by averaging the total number of words in the passage over the total number of gaps. The results indicated that there was an average of 6.35, 7.06 and 7.4 words as contextual support for each gap at the beginning, intermediate and advanced levels respectively. The learners were required to read the passage and choose the best answer fitting with the gaps from the choices given for each blank. One mark was allocated to each correct answer, with a total of 40-50 marks for this task.

The test at each proficiency level was piloted three times, each time with a different group of L1 Persian EFL learners chosen from the language school, but not from the participant pool for the main study. During each pilot, they completed the test followed by a retrospective report where they were asked about their overall viewpoints on the test, test items, and content of the test. Then their tests were scored and an item analysis was conducted to examine the contribution that each test item made to the test. Finally, based on results of the item analysis and the participants' reports on the test, the poor items were identified and revised until no further ones were identified.

During the first pilot, the test at the beginning, intermediate and advanced levels was administered to 9, 7 and 10 participants respectively. As they had prior experience with cloze reading tests, they raised no questions about the test during the exam session. They reported that they had not been familiar with the content of the test and that some of the test items had been too easy and some too difficult. They said that the main problem with the overly difficult items was that more than one answer could have been correct. And the main problem with the too easy ones was that they had been able to dismiss the distracters very easily. The results of an analysis indicated a range of 15, 21 and 17 poor items which had been either too easy or too difficult at the beginning, intermediate, and advanced levels respectively. To determine the poor items, the facility value (e.g., Alderson, Clapham, & Wall, 1995) for each item was computed. This was used to measure the difficulty of an item, and was obtained by dividing the participants who scored correctly by the total number of participants. The cut-off levels used on the item analysis to determine these poor items were .22-.77; .14-.71; and .2-.8 for the beginning, intermediate, and advanced levels respectively. These items were revised by replacing the distracters with more appropriate options. More specifically, the distracters for the overly easy items were replaced with more challenging distracters and those for the overly difficult items were replaced with less ambiguous distracters.

To determine how the revised tests would work, they were administered to a group of 9 participants, 3 in each proficiency group, during the second pilot. The participants at each proficiency level reported that they had had no problems with the test. The results of an item analysis indicated that of the revised test items, those at the beginning level had worked well. The facility value index for this test was in a desirable range of .33-.66. However, a few of the revised test items were still not performing well for the tests at the intermediate and advanced levels. They were either too easy or too difficult. These items were revised once more. The distracters for these items were replaced with more appropriate options.

Finally, the revised tests were piloted with a group of 4 intermediate and 3 advanced participants during the third pilot. The participants reported there were no overly difficult or overly easy test items at this stage. The results of the item analysis also indicated that the facility value index for these tests at the intermediate and advanced levels were in a desirable range of .25-.75 and .33-.66 respectively. An estimate of the duration of the exam session was also obtained for the tests. This was the average time for each test to be completed by the participants at each proficiency level. It included 40, 45 and 50 minutes for the cloze test at the beginning, intermediate and advanced levels respectively. These tests were then finalized for use in main study.

2.3 Procedure

After the cloze tests were developed and piloted by the researcher, and their potential problems were identified and removed, the main study was conducted. As indicated in Table Three, this study was carried out in 4 sessions over a week. In the first session, all the participants were initially required to complete a roughly 90-minute general proficiency placement test. In the language proficiency placement test, they completed a listening section first, and then reading and grammar sections. Then the participants were distributed into three proficiency groups based on their scores in this test. Those participants who obtained scores within a range of 1-30, 31-49 and 50-70 were

placed at the beginning, intermediate and advanced levels respectively. Since the cloze tests were different for each proficiency level, they were conducted in three consecutive rounds. In the first round, the participants at the beginning level completed the test in a 25-minute session. In the second round, the participants at the intermediate level did the same in a 30-minute session. In the third round, the participants at the advanced level were allocated 35 minutes to complete their reading cloze tests. Differences in allotted time at each proficiency level were due to the tests which were different in terms of length (number of words and propositions), and readability. These tests were conducted in the language school where the participants had been chosen for the study. Since the cloze tests were in a multiple-choice format, the participants were just required to choose the correct answer and mark it in their answer sheets.

Table 3. Timetable used for the reading cloze test in this study

Session	Cloze Tests	Duration
One	Beginning Level	25 minutes
Two	Intermediate Level	30 minutes
Three	Advanced Level	35 minutes

Once each participant completed their reading cloze tests, they gave their answer sheets to the researcher along with the retrospective report on these measures. Unlike the participants at the advanced level who had completed the test within the given time, the participants at the beginning and intermediate levels reported that the test had been too demanding and they had not been managed to complete the test.

2.4 Data Analysis

After the tests were completed, each participant was assigned an ID code. Then each test was scored by the researcher. One mark was allocated to each correct answer, with a total of 40, 45 and 50 marks for this measure at the beginning, intermediate and advanced levels.

3. Results

The participants in each proficiency group completed a different reading cloze test. Their performance was analyzed in each proficiency group. Descriptive statistics for reading performance were obtained for each proficiency level. The results of this analysis are provided in Table 4.

Table 4. Descriptive statistics for the level specific cloze reading tests

		N	Score Total	Min	Max	Mean	S.D.
Cloze	Beg.	56	26	4	25	10.25	4.30
	Int.	43	25	4	21	12.04	4.24
	Adv.	41	50	8	44	30.00	8.39

Note. Beg. = Beginning; Int. = Intermediate; Adv. = Advanced

As indicated in Table 4, the reading performance includes the scores of one cloze test at each proficiency level. While the beginning and intermediate participants have closer standard deviations for the cloze tests, the advanced participants have a wider standard deviation for this test. This suggests that the cloze test at the advanced level was better able to discriminate among the participants and indicate individual differences by yielding a wider range of scores than for the beginning and intermediate participants.

4. Discussion

This study examined what the optimum length of a reading cloze test would be for three different proficiency levels. The findings of this study suggest that the optimum length of a reading cloze test may vary according to proficiency level. While it could be lengthy (including 50 empty spaces) at the advanced level, a shorter test would

be more appropriate at the beginning and intermediate levels. One explanation could be the cognitive demanding the lengthy cloze test could have placed on beginning and intermediate participants. The participants at the beginning and intermediate levels reported that the cloze test had been very memory demanding because they had to simultaneously maintain the text information and the given options for each test item to arrive at the most appropriate answer. This could be explained by the memory operations involved in conducting this task. As the prior research on reading assessment (e.g., Cain & Oakhill, 2006; Koda, 2005; Alderson, 2000) points out, a cloze test of this type involves the selection of an answer from a choice of usually 3-5 options. Since the selection of the answer requires the inhibition of irrelevant information followed by updating memory for the next test item, additional memory resources could be employed here (Carretti, Cornoldi, De Beni, & Romano, 2005; Gernsbacher, Varner, & Faust, 1990; Morris & Jones, 1990). However, at the advanced level, the participants reported that they could easily distinguish the sequence between the ideas in the text and show that they had understood in a coherent way. One possible explanation could be that, due to higher L2 knowledge, much of the language processing at the advanced level may have been less controlled, less effortful (less capacity-demanding) and more automatic (e.g., N. Ellis, 2001; Schmidt, 2001; Skehan 1989). More specifically, since the majority of the lower level aspects of reading such as processing orthographic, phonological, semantic and syntactical information is operated automatically (i.e., without adding attention to process meaning), and automatic processes are much less cognitively demanding (e.g., Segalowitz, Segalowitz, & Wood, 1998), these participants do not rely very much on cognitive resources in processing text information.

This study provides further support for the prior research by Brown (1998) where he suggests “longer tests tend to be more reliable than shorter ones, but there is a point at which adding more items can have the opposite effect, as fatigue sets in and the students begin to get discouraged or stop taking the test seriously” (p. 18). This study also provides further credit to Nation’s (2009) proposal of 40-50 test items for a cloze test. What the current study adds to the literature is that the optimum length of a cloze test may rely on the proficiency level. It could be lengthy and include up to 50 test items at the higher proficiency levels. In this study, the results of an item analysis indicated a satisfactory internal reliability of the cloze test at the advanced level. Internal reliability as indicated by Cronbach’s Alpha was .872 for the cloze test with 50 test items here. However, to have a reliable cloze test at lower proficiency levels, the test items should be limited to 20-25. The results of an item analysis indicated a satisfactory range for the internal reliability of the cloze tests at the beginning and intermediate levels only when the test items were limited to 20-25. Cronbach’s Alpha was .730 and .814 for the beginning and intermediate levels respectively here.

5. Conclusion

Based on the results of this study, the optimum length of a cloze test may vary according to a proficiency level. At the advanced level, a cloze test could be long enough and include up to 50 test items. At the beginning and intermediate levels, a shorter cloze test including 20-25 test items may yield more reliable estimates. Findings of the current study provide EFL teachers with an estimated range of test items for a cloze test at each proficiency level. It should be noted that this study was carried out in an EFL context. So, further research should be carried out, particularly in an ESL context, to provide further support for the findings of this study.

References

- Abraham, R. G., & Chapelle, C. A. (1992). The meaning of cloze test scores: An item difficulty perspective. *Modern Language Journal*, 76, 468-479. <http://dx.doi.org/10.1111/j.1540-4781.1992.tb05394.x>
- Alderson, J. C. (2000). *Assessing Reading*. Cambridge: Cambridge University Press. <http://dx.doi.org/10.1017/CBO9780511732935>
- Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.
- Beech, J. R. (1997). Assessment of memory and reading. In Beech, J. R., & Singleton, C. (Eds.), *The Psychological Assessment of Reading*. London: Routledge.
- Brown, J. D. (1989). Cloze item difficulty. *JALT Journal*, 10, 15-42.
- Brown, J. D. (1992). What text characteristics predict human performance on cloze test items. *Proceedings of the 3rd Conference on Second Language Research in Japan*, (pp. 1-26). Niigato, Japan: International University of Japan.
- Brown, J. D. (1996). *Testing in language programmes*. Upper Saddle River, NJ: Prentice Hall.
- Brown, J. D. (1998). Cloze tests and optimum test length. *JALT testing and evaluation SIG Newsletter*, 2(1), 18-21.

- Cain, K., & Oakhill, J. (2006). Assessment matters: Issues in the measurement of reading comprehension. *British Journal of Educational Psychology*, 76, 697-708. <http://dx.doi.org/10.1348/000709905X69807>
- Carretti, B., Cornoldi, C., De Beni, R., & Romano, M. (2005). Updating in working memory: A comparison of good and poor comprehenders. *Journal of Experimental Child Psychology*, 91, 45-66. <http://dx.doi.org/10.1016/j.jecp.2005.01.005>
- Cobb, T. (2002). *Web VocabProfile*. Retrieved October, 2009, from <http://www.Lextutor.ca/vp/>
- Ellis, N. C. (2001). Memory for language. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 33-68). Cambridge: Cambridge University Press.
- Gernsbacher, M. M., Varner, K. R., & Faust, M. E. (1990). Investigating individual differences in general comprehension skill. *Journal of Experimental Psychology: Learning Memory and Cognition*, 16, 430-445. <http://dx.doi.org/10.1037/0278-7393.16.3.430>
- Heaton, J. B. (1988). *Writing English Language Tests*. London: Longman.
- Koda, K. (2005). *Insights into Second Language Reading: A cross-linguistic approach*. Cambridge University Press.
- Laufer, B., & Nation, P. (1995). Vocabulary size and use: Lexical richness in L2 written production. *Applied Linguistics*, 16(3), 307-322. <http://dx.doi.org/10.1093/applin/16.3.307>
- Lesley, T., Hansen, C., & Zukowski/Faust, J. (2003). *Placement and Evaluation Package*. Cambridge: Cambridge University Press.
- Madsen, H. S. (1983). *Techniques in Testing*. New York: Oxford University Press.
- Morris, N., & Jones, D. M. (1990). Memory updating in working memory: The role of the central executive. *British Journal of Psychology*, 81, 111-121. <http://dx.doi.org/10.1111/j.2044-8295.1990.tb02349.x>
- Nation, I. S. P. (2009). *Teaching ESL/EFL Reading and Writing*. Routledge, Taylor & Francis.
- Oller, J. W. (1975). Pragmatics mappings. *Lingua*, 35, 333-334. [http://dx.doi.org/10.1016/0024-3841\(75\)90063-7](http://dx.doi.org/10.1016/0024-3841(75)90063-7)
- Oller, J. W. (1979). *Language tests at school: A pragmatic approach*. London: Longman.
- Rand, E. (1978). The effects of test length and scoring method on the precision of cloze test scores. *UCLA Work papers in Teaching English as a Second language*, 12, 62-71.
- Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 3-32). Cambridge: Cambridge University Press.
- Segalowitz, S. J., Segalowitz, N. S., & Wood, A. G. (1998). Assessing the development of automaticity in second language word recognition. *Applied Psycholinguistics*, 19, 53-67. <http://dx.doi.org/10.1017/S0142716400010572>
- Skehan, P. (1989). *Individual differences in second language learning*. London: Edward Arnold.

Appendixes

Appendix A: Cloze Test at the Beginning Level

Changing Ideas about Art

A lot of modern art is not realistic or beautiful. Artists want to show people a (1) way of seeing things. They want to say (2) about the world they live in (3) paintings and sculptures disturb some (4) because they do not (5) or understand (6) the artist is trying to (7) It often takes many (8) before an artist's work is accepted by the (9)

In 1870, Vincent Van Gogh and Claude Monet were painting in France. They and some other painters (10) many ideas about (11) They used strong, (12) colors, and their (13) were not always very (14) Their art was very (15) from what people were (16) with, and few people liked it. The (17) did not make a lot of (18) with their paintings. Van Gogh (19) only one painting while he was (20)

Today, we (21) this group the Impressionists. Van Gogh's and Monet's paintings are in (22) all over the world. The Impressionists are not only (23) , they are also (24) People everywhere buy their (25) in inexpensive posters and (26) When museums have special (27) of Impressionist paintings, people stand in line (28) for hours to (29) in.

Artists' ideas change (30) , but the public's ideas (31) much more slowly. (32) , artists who are (33) popular in their (34) are not well respected. Those who are (35) can become (36) , but they are (37) popular. Most of today's (38) artists will be (39) in 100 years. Only a few of them will be (40) .

- | | | | |
|-------------------|-----------------|----------------|-----------------|
| 1) a) real | b) secret | c) clear | d) different |
| 2) a) information | b) something | c) stories | d) jokes |
| 3) a) Beautiful | b) Expensive | c) Modern | d) Amusing |
| 4) a) people | b) experts | c) officials | d) parents |
| 5) a) watch | b) like | c) indicate | d) follow |
| 6) a) which | b) when | c) what | d) whom |
| 7) a) observe | b) consider | c) improve | d) express |
| 8) a) years | b) ages | c) experiences | d) records |
| 9) a) public | b) government | c) college | d) museum |
| 10) a) described | b) shared | c) planted | d) received |
| 11) a) culture | b) cinema | c) art | d) photography |
| 12) a) modern | b) fair | c) artificial | d) bright |
| 13) a) papers | b) posters | c) pictures | d) positions |
| 14) a) realistic | b) efficient | c) energetic | d) systematic |
| 15) a) separated | b) different | c) far | d) away |
| 16) a) associated | b) content | c) filled | d) familiar |
| 17) a) group | b) swarm | c) crowd | d) bunch |
| 18) a) medals | b) treasures | c) money | d) gold |
| 19) a) gave | b) watched | c) started | d) sold |
| 20) a) famous | b) alive | c) rich | d) young |
| 21) a) provide | b) tell | c) guess | d) call |
| 22) a) companies | b) universities | c) museums | d) exhibitions |
| 23) a) respected | b) helped | c) offered | d) employed |
| 24) a) social | b) popular | c) honest | d) emotional |
| 25) a) plan | b) performance | c) statue | d) work |
| 26) a) prints | b) lines | c) stamps | d) instructions |
| 27) a) ceremonies | b) exhibits | c) semesters | d) hobbies |
| 28) a) carefully | b) skillfully | c) patiently | d) seriously |
| 29) a) take | b) get | c) fall | d) drop |
| 30) a) heavily | b) perfectly | c) quickly | d) basically |
| 31) a) change | b) make | c) influence | d) continue |
| 32) a) Ideally | b) Immediately | c) Recently | d) Generally |
| 33) a) truly | b) silently | c) likely | d) slightly |
| 34) a) studies | b) schools | c) centuries | d) lifetimes |
| 35) a) educated | b) admired | c) welcomed | d) examined |
| 36) a) strong | b) excited | c) rich | d) tired |
| 37) a) often | b) seldom | c) sometimes | d) usually |
| 38) a) new | b) native | c) tall | d) thin |
| 39) a) promoted | b) rejected | c) forgotten | d) supported |
| 40) a) decisive | b) eager | c) surprised | d) immortal |

Appendix B: Cloze Test at the Intermediate Level

Roles in Human Society

Human beings are creatures of society. They take part in a (1) social system which expects them to (2) certain roles. Social scientists (3) that without roles, society could not (4)

To be (5), members of society need to know how others (6) them to act so that they can act, or not act, in those (7). Let us take student life at a (8) as an example. When new students (9), they do not yet know what their (10) roles are. That is, they do not know what their (11), teachers, or (12) want them to do. To help them (13) quickly and correctly, (14) make them attend a(n) (15) program. They learn the expected (16) for college students. We can label this their (17) role.

In addition to the role, social (18) talk about the (19) role. For our (20) students, this means the (21) that each one has about what (22) behavior at a university is. In order to (23), he or she must know or (24) what others' roles are and then (25) his or her own (26) in relationship to them.

When members of a society (27) perceive the rules of that society, and when their subjective roles are (28) to their prescribed roles, they (29) act in the ways that society (30) them to. That is, they do and say what is (31) correct. The actual (32) of a role, with its (33) behavior, is called the (34) role. Our college students, if they (35) similar prescribed and subjective roles, will (36) obey university roles and (37) with their professors as (38) and with their roommates as (39). Their behavior will fall into the (40) of acceptable student behaviour.

Social (41) say that in order to (42) itself and make sure that its (43) perform their roles, society (44) those members to (45) others' behaviour as acceptable or unacceptable.

- | | | | |
|--------------------|-----------------|---------------|-----------------|
| 1) a) comfortable | b) complex | c) correct | d) continuous |
| 2) a) persuade | b) compress | c) support | d) perform |
| 3) a) affirm | b) evaluate | c) generalize | d) collect |
| 4) a) promote | b) collide | c) function | d) combine |
| 5) a) successful | b) important | c) well-known | d) civilized |
| 6) a) assure | b) assist | c) engage | d) expect |
| 7) a) strategies | b) unions | c) ways | d) shares |
| 8) a) home | b) university | c) barrack | d) hospital |
| 9) a) graduate | b) survive | c) observe | d) arrive |
| 10) a) appropriate | b) associative | c) collective | d) speculative |
| 11) a) parents | b) officers | c) roommates | d) professors |
| 12) a) researchers | b) advisors | c) players | d) volunteers |
| 13) a) construct | b) travel | c) adapt | d) tolerate |
| 14) a) authorities | b) governments | c) carpenters | d) adherents |
| 15) a) celebration | b) orientation | c) formation | d) repetition |
| 16) a) issue | b) introduction | c) schedule | d) behavior |
| 17) a) memorized | b) reinforced | c) migrated | d) prescribed |
| 18) a) researchers | b) scientists | c) officials | d) instructors |
| 19) a) subjective | b) corrective | c) transitive | d) selective |
| 20) a) native | b) sophomore | c) college | d) junior |
| 21) a) inventions | b) statistics | c) documents | d) expectations |
| 22) a) functional | b) cultural | c) suitable | d) relative |
| 23) a) conduct | b) achieve | c) calculate | d) compound |
| 24) a) tie down | b) find out | c) climb up | d) run into |

- | | | | |
|--------------------|-----------------|------------------|----------------|
| 25) a) make up | b) kneel down | c) take out | d) look at |
| 26) a) perceptions | b) conditions | c) distinctions | d) traditions |
| 27) a) wildly | b) clearly | c) occasionally | d) heavily |
| 28) a) related | b) complicated | c) similar | d) significant |
| 29) a) politely | b) recently | c) formerly | d) normally |
| 30) a) relieves | b) lifts | c) expects | d) amuses |
| 31) a) considered | b) designed | c) elaborated | d) included |
| 32) a) opportunity | b) performance | c) influence | d) perspective |
| 33) a) ultimate | b) specific | c) obvious | d) unusual |
| 34) a) weakened | b) internalized | c) enacted | d) decisive |
| 35) a) consult | b) review | c) surprise | d) possess |
| 36) a) probably | b) largely | c) really | d) accurately |
| 37) a) explain | b) interact | c) consume | d) collect |
| 38) a) farmers | b) b) tutors | c) students | d) lecturers |
| 39) a) kids | b) classmates | c) actors | d) friends |
| 40) a) hole | b) range | c) circle | d) space |
| 41) a) experts | b) adherents | c) creators | d) candidates |
| 42) a) compute | b) persuade | c) protect | d) perceive |
| 43) a) resources | b) opponents | c) advisors | d) members |
| 44) a) identifies | b) encourages | c) characterizes | d) criticizes |
| 45) a) certify | b) organize | c) observe | d) judge |

Appendix C: Cloze Test at the Advanced Level

Human Adaptation to Space

“It is important for the human race to spread out into space for the survival of the species,” said world-renowned astrophysicist Steven Hawking. He is far from being alone in his (1) of humans learning to live in places (2) on Earth. A Space Odyssey (3) the possibility of (4) human life in (5) space, and presented a very realistic (6) of spaceflight. Since astronaut Yuri Gagarin (7) the first man to (8) in space in 1961, (9) have researched what (10) are like beyond Earth’s atmosphere, and what (11) space travel has on the human (12).

Although most astronauts do not (13) more than a few months in space, many experience (14) problems when they (15) Earth. Some of these (16) are short-lived: others may be (17).

More than two-thirds of all astronauts suffer from (18) sickness while traveling in space. In the (19) - free environment, the body cannot (20) up from down. The body’s internal (21) system sends (22) signals to the brain, which can result in (23) lasting as long as a few days. A body that is (24) gravity also experiences changes in the (25) of bodily fluids. More fluid than normal (26) up in the face, neck, and chest, resulting in a (27) face, bulging neck (28), and a slightly (29) heart.

Throughout the (30) of a mission, astronauts’ bodies (31) some potentially dangerous (32). One of the most common is (33) of muscle mass and bone (34). Another effect of the (35) environment is that astronauts (36) not to use the muscles they (37) on in a gravity environment, so the muscles (38) atrophy. This, combined with the (39) of fluid to the (40) body and the resulting loss of essential (41) such as calcium, causes bones to (42). Bone density can (43) at a rate of one to two percent a month and, as a result, many astronauts are unable to walk (44) for a few days upon their return to Earth. Exposure to (45) is another serious (46) that astronauts face. Without the Earth’s (47) to protect them, astronauts can be exposed to (48) radiations from the sun and other (49) bodies, leaving them at risk of (50).

- | | | | |
|-----------------------|--------------------|------------------|-------------------|
| 1) a) vision | b) vote | c) vessel | d) venture |
| 2) a) away from | b) similar to | c) next to | d) other than |
| 3) a) established | b) stimulated | c) explored | d) exacerbated |
| 4) a) creating | b) sustaining | b) exclaiming | d) exploiting |
| 5) a) lighter | b) upper | c) larger | d) outer |
| 6) a) theme | b) viewpoint | c) portrayal | d) content |
| 7) a) promoted | b) become | c) submitted | d) negotiated |
| 8) a) travel | b) mediate | c) relocate | d) unite |
| 9) a) physicians | b) psychologists | c) scientists | d) archaeologists |
| 10) a) surroundings | b) arguments | c) certificates | d) conditions |
| 11) a) considerations | b) effects | c) circumstances | d) results |
| 12) a) body | b) personality | c) brain | d) heart |
| 13) a) concentrate | b) spend | c) sleep | d) accept |
| 14) a) analytical | b) celestial | c) theoretical | d) psychological |
| 15) a) go around | b) walk into | c) return to | d) reside in |
| 16) a) ailments | b) advantages | c) adjustments | d) adventures |
| 17) a) taking | b) changing | c) staying | d) long-lasting |
| 18) a) rigid | d) motion | c) modern | d) emotional |
| 19) a) pressure | b) oxygen | c) gravity | d) human |
| 20) a) distinguish | b) encapsulate | c) ventilate | d) enhance |
| 21) a) blood | b) bladder | c) balance | d) breathing |
| 22) a) confusing | b) discriminating | c) circulating | d) coinciding |
| 23) a) battle | b) nausea | c) diamond | d) funeral |
| 24) a) supplied with | b) identified with | c) supported by | d) deprived of |
| 25) a) deduction | b) destination | c) distribution | d) declination |
| 26) a) ends | b) sets | c) backs | d) draws |
| 27) a) choppy | c) puffy | c) catchy | d) cagey |
| 28) a) eyes | b) cheeks | c) laps | d) veins |
| 29) a) endured | b) enlarged | c) enriched | d) engaged |
| 30) a) duration | b) extension | c) formation | d) computation |
| 31) a) accommodate | b) behold | c) experience | d) illustrate |
| 32) a) situations | b) circumstances | c) consequences | d) disorders |
| 33) a) loss | b) injury | c) mixture | d) excess |
| 34) a) severity | b) debility | c) security | d) density |
| 35) a) defenceless | b) featureless | c) weightless | d) tactless |
| 36) a) tend | b) like | c) allow | d) order |
| 37) a) drag | b) rely | c) lay | d) touch |
| 38) a) unlikely | b) consistently | c) surprisingly | d) gradually |
| 39) a) shift | b) generation | c) risk | d) promotion |
| 40) a) controversial | b) imaginary | c) upper | d) supplementary |
| 41) a) elements | b) medicines | c) matters | d) minerals |
| 42) a) reinforce | b) weaken | c) reconstruct | d) expand |

- | | | | |
|--------------------|---------------|-------------------|----------------|
| 43) a) decrease | b) confront | c) mediate | d) amplify |
| 44) a) steadily | b) hesitantly | c) properly | d) unwillingly |
| 45) a) radiations | b) chemicals | c) assaults | d) vacuums |
| 46) a) opportunity | b) hazard | c) characteristic | d) mediocrity |
| 47) a) gravity | b) wideness | c) atmosphere | d) balance |
| 48) a) curious | b) inspired | c) furnished | d) intense |
| 49) a) galactic | b) symbolic | c) telescopic | d) traumatic |
| 50) a) extinction | b) cancer | c) isolation | d) catastrophe |