

How Reading Comprehension Test Diagnoses Vocational College Students' English Reading Competence

Abstract

This paper attempts to explore the possibility of employing reading comprehension test to diagnose vocational college students' strengths and weaknesses in English reading. Conducted with 94 second grade five-year junior college students, this research identifies that the students encountered more difficulties in determining correct discrete grammatical function words than in grasping general meaning of broad scope reading. Besides, without the requirement of reading for main ideas and making inference, the students performed better in demonstrating their reading comprehension ability in terms of reading for details and using reference in the reading text. This research employs both arithmetical and statistical techniques to evaluate the students' performance in the reading comprehension test. It confirms that effective English reading tests can pinpoint students' strengths and weakness from the highly complex characteristics of language skills, and the findings can be used by teachers either to remedy their defects in teaching or to motivate their students to work harder.

Key words:

English reading comprehension, language proficiency assessment, teaching effectiveness

Introduction

Action research can provide many benefits to English teachers. As Grabe and Stoller (2002) suggest, "action research helps teachers develop professionally through the systematic collection and analysis of relevant data and, then, the use of results as the basis for decisions about further professional actions". In order to analyze the complicated data gather through this action research, I've decided to employ Bachman and Palmers' (1996) 'task characteristics framework' as the guideline for my reading comprehension test design. In this research, I've selected an interesting newspaper article that reviews research on the connection between sleep and intelligence. Owing to the commonality of this topic to every test-taker, there should be no influence resulting from their culture differences or prior accumulated knowledge. In order to verify the reliability of my test design and its outcome, all of the students are also assigned a reference score. Calculation of this reference score is based on each student's overall performance in my English reading and vocabulary course in the previous semester. Taking their performance in several quizzes, midterm, final and course assignment altogether into consideration, the trustworthy reference score can be used as an index of the students' English competence. By utilizing both of arithmetical and statistical

techniques, this research aims to evaluate the interrelationship between my test design and its outcome, and consequently gain more insights into the domain of English reading comprehension testing.

Test design in this research

1. Participants

All of the test-takers were from my “Intermediate English Reading and Vocabulary” course in a private vocational college in northern Taiwan. The reason why I chose them as my research subjects was because I was responsible for teaching the same course to both Class A and Class B, totally 94 students (48 and 46 students respectively). Based on my in-class observation, English reading ability of Class B students outperformed Class A students significantly. Consequently, enough samples for analysis could be obtained and meaningful comparison could be made without any difficulty. Furthermore, under the same instruction of reading skills and strategy, variables resulted from their previous reading habits would be reduced. Most important of all, the research results I conclude here will be employed in the practical teaching settings, as a tool to enhance the quality of my English reading teaching.

2. Structure and time allotment

As Hughes (1989) suggests, test designers’ aim of constructing test tasks must be to write items which will permit highly reliable scoring. In order to achieve this goal, the reading comprehension test in this research was designed by me and it consisted of three sections: Section 1 – Selective deletion gap-filling test, Section 2 – Multiple-choice questions, and Section 3 – Ordering task. The test tasks in Section 1 were designed based on the first two paragraphs of the original text. In order to eliminate the influences resulting from the repetition of the same text content distributed to test-takers later, Section 1 was printed separately in exam paper A, and both Section 2 and 3 were printed in exam paper B1 or B2. All of the test-takers were required to accomplish the tasks in either exam paper A plus B1 or A plus B2. Owing to the difference of students’ individual English competence, I intentionally allocated the exam paper B1 and B2 randomly in both classes. By taking this approach, possible bias may be eliminated materially. The whole test was conducted in two stages in 40 minutes: in stage one all of the students were required to complete exam paper A in 15 minutes, and in stage two, they should finish the questions of exam paper B1 or B2 in 25 minutes.

3. Scoring method

The ubiquity of the gap-filling and the multiple-choice techniques reveal the importance of convenience and efficiency to many English teachers. In this reading comprehension test, all of the gap-filling and the multiple-choice questions follow the same principle. For each question, there were one stem and four answer choices. Among the four options, only one

choice was correct. In some cases, there were two incorrect distracters, one almost correct choice, and one correct choice (Broukal, 1997). The scoring criterion for correctness in these two types of questions was straightforward: scores were given to students only when they chose the correct answer. The scoring method for the third section – ordering task was problematic. Since the employment of gap-filling and multiple-choice techniques was timesaving for many English teachers when marking exam papers, as some researchers argued, it may be at the expense of validity (Alderson, 2000). In order to ensure the validity of the third section of the reading test, I employed two kinds of criterion to assess the correctness of test-takers' responses. Method 1 was to mark students' answers according to the common practice: either wholly right or wholly wrong. The students with wholly right answer were scored 100, however, with wholly or partially wrong response were scored 0. Although this method was easy and convenient for the teachers, as Alderson *et al.* (1995) criticize, "the amount of effort involved in both constructing and in answering the item may not be considered to be worth it...." To compensate for this unfairness, I introduced another scoring method that was more complex. Method 2 evaluates students' responses based on both wholly and partially correct answers. For instance, if a student had wholly correct answer, the score was 100. If the student had two elements out of five in correct sequence, 40 scores were given. In case the student had three elements out of five in correct order, the student was scored 60 out of 100.

Test results and discussions

Test results of Class A and Class B are shown in Table 1. Obviously the data indicated that the scores achieved by the students from both classes in each section are significantly below the reference scores. Although the reading text selected in this research was originally designed for low-intermediate to intermediate learners, apparently it was still beyond the test-takers' ability to handle it satisfactorily.

Findings derived from the reading comprehension test are discussed as follows:

Section 1: Broad scope test question v. s. narrow scope test question

In this research, broad scope questions were designed to test students' reading ability in terms of content words, and narrow scope questions were aimed to investigate their competence in identifying correct forms of grammatical function words. As test result shows, both Class A and Class B students performed much better in broad scope questions than in narrow scope questions. The findings reveal the fact that when being under an English reading comprehension test, my students encounter more difficulties in determining correct discrete grammatical function words than in grasping general meaning of broad scope reading.

Table 1

Average Score of Each Section Achieved by Class A and Class B

Test-takers and average scores		Class A	Class B
Sections of the test			
Reference score (100%)		58	66
Section 1.	Broad scope – content words (50%)	22	23
	Narrow scope – function words (50%)	14	14
	Total (100%)	35	37
Section 2.	Test B1 (discrete-point testing approach, 100%)	38	53
	Test B2 (integrative testing approach, 100%)	29	37
	Mixed of Test B1 and B2 (100%)	34	45
Section 3.	Method 1 (scoring criterion is either wholly right or wholly wrong, 100%)	0	2
	Method 2 (scoring criterion is both wholly and partially right, 100%)	38	26

Section 2: Discrete-point test question v. s. integrative test question

The rationale of the test task design in Test B1 and Test B2 originated from two different viewpoints of linguistic reasoning approach: discrete-point approach and integrative approach. Test results indicate that both groups of students gained higher scores in Test B1 than in Test B2. Furthermore, test-takers' performance in Test B1 was even the best among the three sections of the test. The findings confirm that the discrete-point test questions are easier for the test-takers to fulfill. Without the requirement of reading for main ideas and making inference, test-takers perform better in demonstrating their reading comprehension ability in terms of reading for details and using reference in the reading text.

Section 3: A re-examination of integrative test task

This section was designed based on the indirect relationship testing approach. Students can complete the assigned test task only when they have fully understanding of the whole text. Students' test result of Section 3 is rather disappointing: only 1 out of 94 test-takers could make precise prediction to the order of the five given sentences. As the result, this should be the most difficult part of the whole reading comprehension test. Although the test result is absolutely unsatisfactory to any test constructor, the two kinds of scoring methods employed in this research are still worth of noticing. By employing Method 1, the average scores of Class A and Class B are 0 and 2 respectively. But by employing Method 2, the scores rise up to 38 and 26. Comparing with the other average scores in the Section 1 and 2, the scores calculated by Method 2 are relatively more meaningful for research analysis. In other words,

when taking all of students' possible efforts involved in constructing and fulfilling the ordering task into consideration, students' real language competence can be identified more effectively.

In addition, correlation matrices for tests result of Class A and B can also be utilized to diagnose students' performance in each section of the test. Table 2 shows that the total score of Class A correlates with Section 1 at 0.4562, with Section 2 at 0.4951, and with Section 3 at 0.8895. Obviously, students in Class A are skilled in fulfilling Section 3 ordering task, and have relatively insufficient ability when taking the tests in forms of gap-filling or multiple-choice questions. However, when making the same comparison with the figures in Table 3, it is clear that students in Class B perform almost equally well in the three sections of the test. Although total score correlates with Section 2 at the highest (0.7726), the correlations between the total score and Section 1 or Section 3 still show the high commonality. The findings further reveal the fact that when taking a reading test, students with higher level of language competence can obtain their scores equally from every section of the whole test. However, students with lower level of language competence might perform significantly better in certain section, and the scores they achieved there might take up a big part of their total scores.

Table 2

The Correlation Matrices for Tests Result of Class A

Correlations	Ref. score	Total	Section 1	Section 2	Section 3
Ref. score	1.0000	0.2652	0.3160	0.0630	0.1805
Total		1.0000	0.4562	0.4951	0.8895
Section 1			1.0000	-0.0552	0.1824
Section 2				1.0000	0.2108
Section 3					1.0000

Table 3

The Correlation Matrices for Tests Result of Class B

Correlations	Ref. score	Total	Section 1	Section 2	Section 3
Ref. score	1.0000	0.4240	0.4740	0.5046	0.0165
Total		1.0000	0.6172	0.7726	0.6999
Section 1			1.0000	0.4637	0.0213
Section 2				1.0000	0.2402
Section 3					1.0000

Conclusion

From my viewpoint, English reading comprehension tests should not be used by vocational college teachers as a means of threat to force students to study. Instead it should be employed as a tool to help students diagnose their strengths and weaknesses in their process of learning.

As discussed in this research, higher English competence students may perform equally well in every part of the test, but lower English competence students may only be good at certain types of questions. Since the lower competence students are the ones demanding more aids and encouragement, teachers can employ various techniques of testing to accomplish these tasks. Effective English reading tests can pinpoint students' strengths and weaknesses from the highly complex characteristics of language skills, and the findings can be employed by teachers either to remedy their defects in teaching or to motivate their students to work harder. For instance, when teachers perceive that their students perform unsatisfactory in the aspects of vocabulary or grammar, or have insufficient readings skills of grasping the main idea or making inference, they can strengthen their teaching effectiveness by providing remedial courses based on the information they gathered from the reading tests. By fulfilling these missions, theories of English reading comprehension test can be turned from a tool of assessment into practical use of strengthening teaching effectiveness.

References

- Alderson, J. C. (2000). *Assessing Reading*. Cambridge, England: Cambridge University Press.
- Alderson, J. C., Clapham, C., & Wall, D. (1995). *Language Test Construction and Evaluation*. Cambridge, England: Cambridge University Press.
- Bachman, L. F. & Palmer, A. S. (1996). *Language Testing in Practice*. Oxford, England: Oxford University Press.
- Broukal, M. (2002). *In-A-Flash: Reading for the TOEFL*. NJ: Thomson Learning.
- Grabe, W. & Stoller, F. L. (2002). *Teaching and Researching Reading*. Harlow: Pearson Education.
- Hughes, A. (1989). *Testing for Language Teachers*. Cambridge, England: Cambridge University Press.