Validity Study on the Test Paper of the Academic English Reading and Writing Course

GUO Sisi, CHEN Zheng
University of Shanghai for Science and Technology, Shanghai, China

The aim of this study is to conduct the validity analysis of the test paper of the academic English reading and writing course referring to the freshmen of non-English majors in a university of Technology in Shanghai. The validity is mainly analyzed from these three aspects: the content and construct validity, the reaction process, and the internal structure of the test paper. And the principles of language assessment in the design of the test paper are given in this study. The validity study was also conducted in a class with 42 freshmen, and the scores of them were analyzed through the internal correlation coefficient from Statistic Package for Social Science (SPSS) 25.0. Finally, the validity of the test paper is concluded.

Keywords: the test paper of the academic English reading and writing course, validity, the internal correlation coefficient

Introduction

In recent years, academic English is continually developing, and it has had its own research field. The academic English teaching is to help college students understand the basic model and basic language norms of the academic English. And this course is a bridge between the basic English and the academic English. And researchers have paid more attention to the quality of the academic English teaching.

Brief Introduction of the Test Paper of the Academic English Reading and Writing Course

The test paper of the academic English reading and writing course is mainly analyzed in this paper to see if it can test students’ academic English reading and writing ability. And the test paper is from a class with 42 freshmen of non-English majors in a university of Technology in Shanghai.

The test paper consists of five parts. The first three parts are objective subjects, that is, academic vocabulary, reading comprehension, and academic knowledge; the latter two are subjective subjects, that is, summary and academic writing. And each part of the test paper will be well introduced in the content and construct validity.

GUO Sisi, Master, student, School of Foreign Languages, University of Shanghai for Science and Technology, Shanghai, China.

CHEN Zheng, Dr., associate professor, School of Foreign Languages, University of Shanghai for Science and Technology, Shanghai, China.
Validity Analysis of the Test Paper

The Content and Construct Validity

Content and construct are two key aspects of validity. Academic vocabulary, the first part of the test paper, has 30 items and each for 0.5 score. The test scope of this part is among the core academic words from each unit of the academic book taught by the teacher during the semester, which can help students clearly and comprehensively master the core academic vocabulary they have learned. The second part is reading comprehension, which accounts for 30% of the whole test paper, two scores for each item, aiming to test students’ ability of understanding and analyzing the passage, which is most important for academic writing. The third part is academic knowledge. Students need to choose A (True) or B (False). This part accounts for the least, only 5%, but it also matters, mainly testing students’ knowledge of the academic English norms, which helps students form a rigorous academic thinking. The fourth and fifth parts are summary and academic writing, the two most important parts in the test paper, which account for 50%. In the fourth part, students need to read an academic article about 500 words, and then answer two questions and write a summary based on it. This part is to test students’ ability to extract the general idea of the article. The last part is academic writing, which is closely related to the fourth part including students’ observations and students’ opinions. This part can foster students’ thinking of the academic writing, greatly contributing to their later academic research.

Comprehensively, students are tested from the easier parts to the more difficult parts and the last two parts—summary and academic writing are based on the first three parts. And students’ academic words, academic knowledge, ability of understanding academic passages and ability of academic writing are all properly tested in this test paper.

Reaction Process

The reaction process is also a main source of validity. The reaction process consists of two parts. The first is students’ reaction to the tasks of the test paper; the other is how recorders react when they record subjective parts.

As for students’ reaction to the tasks of the test paper, the think aloud method can be used. Let students say their own thinking while they are taking the test and see if the thinking process is the same as that the teacher has taught in the class. If they are not the same, students’ reaction can be further analyzed from the perspective of effectiveness. And if the outcome is that the students’ reaction is effective, then test makers should reform the test paper accordingly. The method of the interview or the questionnaire survey can be also employed among students immediately after the test to know students’ thinking while they are taking the test. The method of the interview can be also employed among recorders to see how students are recorded and whether the recorders follow certain recording criteria.

The Internal Structure of the Test Paper

The validity of the test paper is also closely related to the internal structure of it. And there are five parts in this test paper. Basically, it is designed from the easier parts to the more difficult parts. And the academic vocabulary, academic reading comprehension, and academic knowledge parts lay the foundation for the latter two parts. In other words, only students master the basic knowledge of the academic English, can they be able to start the academic writing, so to some extent each part of this test paper can examine the corresponding ability of students.
Principles in the Design of the Test Paper

In the process of the design of the test paper, the principle—target language use domain proposed by Professor Bachman and Palmer is used. The purpose of the test is to test students’ ability of the academic English. And the Assessment Use Argument (AUA) modal proposed by Processor Bachman is a key theory used in the design of the test paper (Xu, 2012). The test consequences are beneficial to stakeholders. For instance, according to the test consequences, students can know which part of the course they have not mastered; as for test makers, they can see whether the test is proper to test students’ ability of the academic English reading and writing. Besides, every part of the test paper is fair to every student and do not favor certain group. What’s more, the test score can basically represent each student’s academic English reading and writing level. As for the scoring methods, the objective parts are scored by computers, while the subjective parts are scored according to the concrete scoring criteria by six teachers.

Data Analysis

The data are from the scores of 42 freshmen in one class of non-English majors in a university of Technology in Shanghai. The internal correlation coefficient can be used to test the construct validity of the test paper. Therefore, the test scores of students have been analyzed by SPSS 25.0. The consequences have been analyzed as follow.

Table 1
Correlations Between Each Part of the Test Paper

|                      | Academic vocabulary score | Reading comprehension score | Academic knowledge score | Summary and academic writing score |
|----------------------|---------------------------|----------------------------|--------------------------|-----------------------------------|
| Academic vocabulary  |                           |                            |                          |                                   |
| score                | Pearson correlation       | 1                          | 0.154                    | 0.235                             |
| Sig. (2-tailed)      |                           | 0.329                      | 0.134                    | 0.052                             |
| N                    |                           | 42                         | 42                       | 42                                |
| Reading comprehension|                           |                            |                          |                                   |
| score                | Pearson correlation       | 0.154                      | 1                        | 0.281                             |
| Sig. (2-tailed)      |                           | 0.329                      | 0.072                    | 0.046                             |
| N                    |                           | 42                         | 42                       | 42                                |
| Academic knowledge   |                           |                            |                          |                                   |
| score                | Pearson correlation       | 0.235                      | 0.281                    | 1                                 |
| Sig. (2-tailed)      |                           | 0.134                      | 0.072                    | 0.036                             |
| N                    |                           | 42                         | 42                       | 42                                |
| Summary and academic |                           |                            |                          |                                   |
| writing score        | Pearson correlation       | 0.302                      | 0.310                    | 0.324                             |
| Sig. (2-tailed)      |                           | 0.052                      | 0.046                    | 0.036                             |
| N                    |                           | 42                         | 42                       | 42                                |

Note. *Correlation is significant at the 0.05 level (2-tailed).

The relationship between every two parts of the test paper was investigated using a Pearson product-moment correlation coefficient. Only the reading comprehension score and the summary and academic writing score, the academic knowledge score and the summary and academic writing score have a correlation but the correlation is small ($r = 0.310$, $N = 42$, $p < 0.05$. $R^2 = 9.6\%$; $r = 0.324$, $N = 42$, $p < 0.05$. $R^2 = 10.5\%$ respectively). Therefore, the correlation between different parts of the test paper is low.
The relationship between the academic vocabulary score and the total score of every student was investigated using a Pearson product-moment correlation coefficient. There is a middle correlation between the two variables: $r = 0.602$, $N = 42$, $p < 0.01$. $R^2 = 36.2\%$.

The relationship between the reading comprehension score and the total score of every student was investigated using a Pearson product-moment correlation coefficient. There is a high correlation between the two variables: $r = 0.725$, $N = 42$, $p < 0.01$. $R^2 = 52.6\%$.

The relationship between the academic knowledge score and the total score of every student was investigated using a Pearson product-moment correlation coefficient. There is a middle correlation between the two variables: $r = 0.538$, $N = 42$, $p < 0.01$. $R^2 = 29.1\%$. 

### Table 2

**Correlations Between the Academic Vocabulary Score and the Total Score of Every Student**

| Correlations                      | Academic vocabulary score | The total score of every student |
|----------------------------------|---------------------------|---------------------------------|
| Pearson correlation              | 1                         | 0.602**                         |
| Sig. (2-tailed)                  |                           | 0.000                           |
| $N$                              | 42                        | 42                              |
| Pearson correlation              | 0.602**                   | 1                               |
| Sig. (2-tailed)                  | 0.000                     |                                 |
| $N$                              | 42                        | 42                              |

**Note.** ** Correlation is significant at the 0.01 level (2-tailed).**

### Table 3

**Correlations Between the Reading Comprehension Score and the Total Score of Every Student**

| Correlations                      | Reading comprehension score | The total score of every student |
|----------------------------------|-----------------------------|---------------------------------|
| Pearson correlation              | 1                           | 0.725**                         |
| Sig. (2-tailed)                  |                            | 0.000                           |
| $N$                              | 42                          | 42                              |
| Pearson correlation              | 0.725**                     | 1                               |
| Sig. (2-tailed)                  | 0.000                       |                                 |
| $N$                              | 42                          | 42                              |

**Note.** ** Correlation is significant at the 0.01 level (2-tailed).**

### Table 4

**Correlations Between the Academic Knowledge Score and the Total Score of Every Student**

| Correlations                      | Academic knowledge score | The total score of every student |
|----------------------------------|--------------------------|---------------------------------|
| Pearson correlation              | 1                        | 0.538**                         |
| Sig. (2-tailed)                  |                          | 0.000                           |
| $N$                              | 42                       | 42                              |
| Pearson correlation              | 0.538**                  | 1                               |
| Sig. (2-tailed)                  | 0.000                    |                                 |
| $N$                              | 42                       | 42                              |

**Note.** ** Correlation is significant at the 0.01 level (2-tailed).**
The relationship between the academic knowledge score and the total score of every student was investigated using a Pearson product-moment correlation coefficient. There is a middle correlation between the two variables: $r = 0.538$, $N = 42$, $p < 0.01$. $R^2 = 28.94\%$.

**Table 5**

| Correlations | Summary and academic writing score | The total score of every student |
|--------------|-----------------------------------|----------------------------------|
| Pearson correlation | 1 | 0.765** |
| Sig. (2-tailed) | - | 0.000 |
| $N$ | 42 | 42 |
| Pearson correlation | 0.765** | 1 |
| Sig. (2-tailed) | 0.000 | - |
| $N$ | 42 | 42 |

**Note.** **Correlation is significant at the 0.01 level (2-tailed).**

The relationship between the summary and academic writing score and the total score of every student was investigated using a Pearson product-moment correlation coefficient. There is a high correlation between the two variables: $r = 0.765$, $N = 42$, $p < 0.01$. $R^2 = 58.5\%$.

Comprehensively, the correlation between different parts is low but each component has a middle or high correlation coefficient with the total score. Therefore, the internal correlation of the test paper is middle to high. So, the construct validity of the test paper is middle to high in terms of the internal correlation.

**Conclusion**

To sum up, the test paper has a moderate to high validity. In a degree, the content and the internal structure of the test paper are relatively proper. However, there are some shortcomings in the study. Firstly, due to the lack of qualitative research, students’ reaction during the test and recorders’ reaction during recording the test paper cannot be known. Secondly, since the scores of each item of each part of the test paper cannot be collected, the internal correlation coefficient of each item of each part of the test paper cannot be known. Therefore, in the future research, the data should be fully collected and the quantitative research and the qualitative research should be mixed, and then the validity of this test paper will certainly be more comprehensively analyzed.

**References**

Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.

Bachman, L. F. (2004). *Statistical analyses for language assessment*. Cambridge: Cambridge University Press.

Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice*. Oxford: Oxford University Press.

Cai, J. G. (2017). The challenges and subversion of foreign language teaching ideas in colleges and universities: Take *Guide to College English Teaching* as an example. *Foreign Language Education*, (1), 6-10.

Chen, Z. (2018). The teaching research of the academic English courses in the university of technology. *English Teachers*, (1), 10-11.

Xu, Q. L. (2012). The evolution and development of the language evaluation effectiveness theory—An inquiry based on the AUA Framework. *Journal of Xi’an International Studies University*, (3), 20-24.

Zou, S. (2012). *Language testing*. Shanghai: Shanghai Foreign Language Education Press.