What is the Difficulty level of Question Types in the IELTS Academic Reading Sub-test

Ruwen Zhang
Faculty of Education and Social Work
University of Auckland
New Zealand

John Hope
Faculty of Education and Social Work
University of Auckland
New Zealand

Abstract

A high grade IELTS (International English Language Testing System) result is one of the most acceptable forms of evidence regarding Chinese candidate's language proficiency for entering foreign universities or colleges, so attaining the required score in each band of the IELTS is extremely important for Chinese students. The purpose of our study is to unprecedentedly identify the difficulty level of all nine question types in the IELTS reading sub-test so as to help students to make improvements in their band score. Our study utilised a framework named task characteristics and via a mixed methodology, we could determine the difficulty level in the academic reading sub-test of the IELTS. The findings reveal that the Chinese student test participants think the question type of Gapfilling is the easiest, while the question type of Matching--finding correct endings is the most difficult. The findings also found that the difficulty level of the question types could have an impact on Chinese students' performance in the IELTS test. The significance of our study is that it could help Chinese IELTS reading teachers to adjust their teaching plans to better prepare their students for the IELTS reading sub-test. They could also be able to more accurately predict their students' performance according to the combination of different question types in IELTS reading tests.

Keywords: Chinese students, academic reading sub-test in the IELTS, difficulty level of question types;

1.0 Introduction

The IELTS, which consists of four sub-tests: listening, reading, writing, and speaking, has been recognised as a trustworthy English language test worldwide (Yang & Badger, 2014). It examines the language proficiency of people who want to study or work in environments where English is used as a language of communication (Is'haaq & Seyyed, 2013). It is thus set as one of the English language entry requirements by many overseas universities, colleges, and institutions, leading to numerous students choosing to enrol in the IELTS test every year. For example, in 2014, 2.2 million candidates across the world sat an IELTS test (IELTS, n.d.). Currently, there is an increasing number of Chinese students studying abroad to get a foreign university qualification (Luo, 2017), which outnumbers any other countries in the world (Li et al., 2017), so achieving the required IELTS score result is a necessary choice for them.

Although many IELTS related research papers have been published, few are directly connected with the difficulty level of question types in the academic reading sub-test, and research publications specifically targeting at Chinese IELTS students are almost nonexistent (IELTS, n.d.). The topic of question types in the reading section, therefore, has not yet been studied thoroughly for Chinese test takers, even though it has been shown to affect test takers' performance (Behfrouz & Nahvi, 2013). In order to help Chinese students to improve their IELTS academic reading band scores further, the main research question of our study was to determine the difficulty level in the IELTS reading sub-test. But before addressing this question, the relating literature must be discussed

2.0 Literature Review

The framework of the test task characteristics used in our research is based on the theoretical model devised by Bachman in 1990 and revised by Bachman and Palmer in 2010. They improved their framework and created a more advanced model that was initially targeted at designing and constructing language tests, but due to its detailed descriptions of tasks, the model also provides a useful paradigm for empirical studies and other research (Carr, 2006). The significance of the test task characteristics outlined is that they are described in a systematic way regarding tests and test procedures (Bachman & Palmer, 1996). The model postulates a series of features for describing five aspects of a test task: setting, test rubric, input, expected response, and the relationship between input and response (Bachman & Palmer, 2010).

Many researchers have used this test task characteristics model to do test related studies (Carr, 2006; Clapham, 1996; Freedle & Kostin, 1999; Perkins, Gupta, & Tammana, 1995), and they mainly focused on the relationship of these features to the test takers' overall performance. Among them, the most relevant study to our research was conducted by Behfrouz and Nahvi (2013). They investigated the effect of task characteristics on Tehran learners' performance in the reading section of the IELTS and confirmed that test characteristics had a significant effect on test takers' performance. They also reported participants performed differently in the reading test when they confronted different question types. Compared with our study, however, Behfrouz and Nahvi only focused on the effect of the task characteristics by using three question types, rather than identifying the difficulty level of all question types.

Among the five Bachman and Palmer's characteristics, our study focused on the aspect of 'input' that "consists of the material contained in a given test task, which the test takers are expected to process in some way and to which they are expected to respond" (Bachman & Palmer, 2010, p. 52). This characteristic is described two perspectives: the format and the language. The format refers to how the input attribute is presented to the test participants, and it includes channel (aural or visual), form (language or non-language), language (native or non-native), length (short or long), type of input (item or prompt), degree of speediness (the rate of processing the test), and vehicle (the means of delivery of the input). The language feature of the input is connected to the nature of the language used in a test, which contains three aspects: the organisational characteristic demonstrates the grammar and textual property, such as cohesion and rhetorical devices; the pragmatic characteristic includes function characteristic and sociolinguistic characteristics; the topical refers to the relevant topics in the input.

Utilising the Bachman and Palmer definitions, the question types in the reading sub-test of the IELTS belong to the 'input' characteristic, and because the Bachman and Palmer study suggests that test task characteristics could have an impact on students' performance, we can draw the conclusion that the question types of the IELTS reading sub-test could affect students' performance. That is to say, identifying the difficulty level of all question types is a significant step towards improving students' performance.

3.0 Methodology

3.1 Settings

Our study was conducted in the IELTS department within a large, private English language teaching institution that has more than 1,000 learning centres across China. It has a very good reputation for teaching quality in China, so most Chinese students would choose this institution to prepare for all kinds of English language tests, including IELTS, TOEFL, and SAT (Scholastic Assessment Test) (New Oriental Education & Technology Group, n.d.; Tan, 2012).

3.2 Participants

The researchers invited all the students and the reading teachers in the IELTS department as participants. Student participants were needed to help establish their perception of the difficulty level of each question type. The reading teachers all understood the reading subtest in depth, as they had received professional training offered by their institution and the IELTS official organisation, so their judgement of the difficulty level was also necessary for our study. Another rationale for inviting the teachers to participate was to eliminate the influence of the difficulty level of the testing passages, as the reading passage also belong to the previously discussed 'input' test task characteristics, which could also have an impact on students' performance. But it would be impossible to determine the difficulty level of the question types if they were detached from the reading passage. In our research, therefore, the teachers' participation was to eliminate the impact of the reading passage themselves on student's performance. To be specific, if the teachers' overall judgment of the difficulty level was similar to that of the students, it suggested that the reading passage themselves were less likely to interfere with the accuracy of our research results.

Through the participation of the students and the teachers, we could determine their joint judgments in terms of the difficulty level of all question types, and by comparing the results from both the students and the teachers, the findings and results would be more objective.

3.3 Procedures

The main study included two stages:

Main study stage 1: Three IELTS academic reading tests In order to include all nine question types, we randomly selected three IELTS academic reading tests from the series of Cambridge English IELTS, which is the official preparation material published by Cambridge University Press. The reason for using this series is "the CUP (Cambridge University Press) materials are made up partly of retired official materials, and so were thought to reflect better than many other commercial materials the actual nature of the official test" (Moore, Morton, & Price, 2012, p. 15). The participates were expected to do all three reading tests one by one, and in total, 170 out of 200 (85%) students volunteered to take part in the three tests. Main study stage 2: Teachers' interview All of the four IELTS reading teachers in the IELTS department voluntarily participated in a focus group interview. The teachers' interview contained nine questions about the IELTS reading teachers' self-perceived difficulty level of all the question types.

3.4 Data Analysis

The analysis of the quantitative data from the students' test results was to calculate the error rate of all the question types in order to identify their difficulty level. We used IBM SPSS 23 (Statistical Product and Service Solutions Version 23)'s base system to determine the extent to which data was missing, assess the excessive repeaters, and then calculate the final result. Data analysis of the qualitative approach was made through coding the field notes from the reading teachers' interviews by hand, which was completed by sorting their answers into different categories of question types. The last step was to compare the results of the two approaches to draw conclusions.

4.0 Results

4.1 Students Quantitative Results

In order to clearly identify the difficulty level of all the nine question types, we calculated the error rate for each type of question across all three tests. There were 59, 61, and 50 participants taking the three reading tests respectively. The tests comprised a total number of 120 question items, and the result of the error rate analysis is as follows

Table 1. The error rate of the nine question types

Question types	Error rate
Matching--finding correct endings	33.2%
Multiple choice--four choices	31.5%
Matching--finding paragraphs	30.4%
List of headings	25.9%
Summary	22.7%
True/False/Not given	22.5%
Multiple choice--more than four choices	21.7%
Classification	21.2%
Gap-filling	20.2%

Although the differences between the percentages was small, three groups could be identified for the ease of discussion. We labelled the question type of *Matching—finding correct endings, Multiple Choice--four choices* and *Matching--finding paragraphs* to be relatively difficult. The question type of *List of headings* was assigned a difficulty level of intermediate. The question types: *Summary, True/False/Not given, Multiple Choice—more than four choices, Classification,* and *Gap-filling* were categorised as comparatively easy for students.

4.2 Teachers Qualitative Results

The perceived difficulty level of question types from the four IELTS reading teachers were almost equivalent to that of the result of students. The four teachers also thought the question types of Matching--finding correct endings, Multiple Choice—four choices and Matching--finding paragraphs were hard for students to deal with; and the question types of List of headings, True/False/Not given, Multiple Choice--more than four choices, Classification, and Gap-filling as easy to answer. But the only difference was they identified the question type of Summary as slightly difficult.

4.3 Summary

In summary, the level of difficulty for most of the question types was consistent between the results of the quantitative data and the qualitative data. To be more explicit, the question type of Matching--finding correct endings was the most difficult; Matching--finding paragraphs was relatively difficult. In addition, the question types of True/False/Not given, Multiple choices--more than four choices, and Classification were comparatively easy, and the question type of Gap filling was classified as the easiest.

But there were different opinions regarding the difficulty level in question types of List of headings and Summary. The quantitative result classified List of headings as intermediately difficult, but the qualitative result listed it as easy. The question type of Summary was considered as an easy question in the quantitative findings,

while the qualitative findings regraded it as intermediately difficult. However, as there was no huge discrepancy in terms of the difficulty level of the two question types, the small differences in the ranking were neglected.

The findings and results do address the main research questions in our study, but they also have some limitations and generate some implications for further research.

5.0 Conclusion and Discussion

Previous literature review has confirmed that test task characteristics could impact students' performance in a test, but the significance of our study is that we proved the question types in the IELTS academic reading test belonged to the test task characteristics of 'input' for the first time, and they will have an influence on students' test outcomes. The final results of our study suggest some implications for both students and teachers in China. For students, after they understand the difficulty level of each question type, they could make a more reasonable study plan. For example, they could spend more time on some easy question types instead of allocating their time equally on all question types. In this way, they could efficiently achieve their required score. For teachers, the most important implication of our study was the explicit suggestions about how to thoughtfully arrange their teaching plan. After identifying the difficulty level of each question type, Chinese IELTS teachers could readjust their teaching plans to specifically address the most difficult question types in their teaching, which should allow them to better help their students. In addition, the results could help them predict the performance of their students in the authentic IELTS academic reading tests more accurately according to the combination of different question types. For example, if a test consists of more high difficulty level question types, most of the students may not get a sound result, and vice versa. It is hoped that these suggestions could be a valuable tool for teachers to serve Chinese students better and help them achieve their desired band score.

The research generated some implications for future development. The first limitation was that our study was conducted in one branch of one institution. Although there are other language training institutions in China, we could not get permission to do the study there, so future research across more than one institution could identify whether these results can be generalised into other Chinese IELTS teaching institutions. Our study also lacks the detailed analysis of the likely influencing factor that affect the determined difficulty level of each question type. Due to the limitation of timespan, we did not go further into the explicit feature of each type of question, which left a larger area for further exploration regarding what factors could actually influence the difficulty level of question types in the IELTS academic reading sub-test.

Finally, our research was intended to serve Chinese students and teachers, so the findings and results may not be valid for students and teachers in other countries. Future research could identify whether the same difficulty level could also apply to students and teachers in other places. With more and more Chinese students going abroad for studying and getting better academic qualification, understanding the unexplored areas of IELTS itself is critical. Our study ex plains the questions types in the IELTS academic reading section, which offers some new ideas for the language teaching in China and gives teachers solid evidence to come up with innovative teaching plans for their students. We hope that our study would start a new chapter to promote more Chinese related research in the field of English language testing and bring more benefits for Chinese students and teachers.

References

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

Bachman, L. F., & Palmer, A. S. (2010). *Language assessment in practice*. Oxford: Oxford University Press.

Behfrouz, B., & Nahvi, E. (2013). The Effect of Task Characteristics on IELTS Reading Performance. *Open Journal of Modern Linguistics, 3(1)*, 30-39.

Carr, N. T. (2006). The factor structure of test task characteristics and examinee performance. *Language Testing, 23,* 269-289.

Clapham, C. (1996). *The development of IELTS: a study of the effect of background knowledge on reading comprehension*. Cambridge: University of Cambridge Local Examinations Syndicate.

Freedle, R., & Kostin, I. (1999). Does the text matter in a multiple-choice test of comprehension? The case for the construct validity of TOEFL's minitalks. *Language Testing*, (*16)*, 2-32.

IELTS. (n.d.). *IELTS.* Retrieved from https://www.ielts.org Is'haaq, A., & Seyyed, M. A. (2013). Comparing the contribution of vocabulary breadth to IELTS and TOEFL reading subtests. *Porta Linguarum, 20,* 135-151.

Li, Z., Heath, M. A., Jackson, A. P., Allen, G. E. K., Fischer, L., & Chan, P. (2017). Acculturation Experiences of Chinese International Students Who Attend American Universities. *Professional Psychology: Research and Practice, 48*(1), 11-21.

Luo, H. (2017). More Chinese students set to study overseas. Retrieved from https://www.telegraph.co.uk/news/world/china-watch/society/more-students-to-study-overseas/

Moore, T., Morton, J., & Price, S. (2012). Construct validity in the IELTS Academic Reading test: A comparison of reading requirements in IELTS test items and in university study. *IELTS Research Reports Volume, (11),* 1-89.

New Oriental Education and Technology Group. (n.d.). *Overview*. Retrieved from http://www.neworiental.org/english/who/201507/8213540.html

Perkins, K., Gupta, L., & Tammana, R. (1995). Predicting item difficulty in a reading comprehension test with an artificial neural net- work. *Language Testing, 12,* 34-53.

Tan, K. (2012). Chinese Education Stock Still Rates an "A". *Barron's, 92*(17), 9.

Yang, Y., & Badger, R. (2015). How IELTS preparation courses support students: IELTS and academic socialisation. *Journal of Further and Higher Education: Promoting student engagement and well-being, 39*(4), 438-465.