



Intelligent Evaluation Methods for English Reading Comprehension Ability and Their Effectiveness Analysis

Le Tian

Chengdu University, Chengdu, Sichuan, China

tianle@cdu.edu.cn

Abstract. The College English Test (CET), comprising the CET-4 and CET-6 exams, serves as a critical tool for assessing English proficiency across the nation and has become an essential benchmark for the enrollment and graduation of university students. Recognizing the significance of reading in the mastery of English, it is evident that the current English reading comprehension skills of the majority of students are insufficient to achieve commendable scores in the CET exams. This paper seeks to delve into the English reading comprehension capabilities of college students. In the era of technological advancement, the integration of intelligent assessment techniques aligns with the evolving educational landscape, offering a promising avenue for enhancing students' reading comprehension skills. By employing experimental and survey methodologies, this study aims to evaluate the efficacy of intelligent assessment methods on improving reading comprehension. The findings from the experimental group indicate a marked improvement in reading comprehension levels, with an enhancement degree ranging from 20% to 30%. In contrast, the control group exhibited a minimal change, with an improvement rate between 3% and 7%. These results underscore the effectiveness of intelligent assessment methods in significantly bolstering students' reading comprehension abilities.

Keywords: CET-4 and CET-6, Reading Comprehension, Intelligent Assessment, Validity Analysis.

1 Introduction

These attributes are indispensable for students to navigate the complexities of career progression and societal demands within an increasingly globalized world. The advent of information technology has catalyzed a transformation in the administrative and evaluative paradigms of standardized assessments.

The paper first recognizes the importance of English language learning and presents the benefits of multiple assessment methods. Secondly, the paper describes other scholars' insights on reading comprehension and strategies for intelligent assessment methods, etc. in the related work section, and then puts forward the argument of this paper. In the third part, the paper provides a theoretical exposition, then discusses intelligent assessment methods, constructs a model, and launches an investigative experiment. In

the fourth part, the paper tests the designed system and conducts group experiments to obtain data results. Finally, the paper is summarized, presenting problems and room for improvement.

2 Related Works

A number of scholars have investigated various theories related to reading comprehension. For example, Paulo Pirozelli et al. aimed to improve reading comprehension skills with datasets related to oceanography, the Brazilian coast, and climate change, focusing on the development and evaluation of Pirà 2.0 [1]. Based on the impact of online oral reading on English language learners, Liqaa Habeb Al Obaydi investigated the effect of his practice on the reading comprehension and anxiety levels of English language learners[2]. Ivan Mardini presented a deep learning system for assessing reading comprehension with the contribution of developing open assessment tools [3]. Nawras Khudhur explored the use of partial decomposition in concept maps to improve reading comprehension and reduce cognitive load [4]. Sophie Gruhn argued that a dynamic approach was needed to understand changes in children's reading interpretations[5]. In Text Comprehension and Reading Learning, Lucia Mason investigated the effects of salience on text comprehension [6]. Anurag Shukl analyzed the role of machine reading comprehension played by DeepInsight[7]. Pengming Wang developed machine reading comprehension models using query reconstruction and deep learning techniques [8]. Samreen Kazi provided a comprehensive survey of deep learning techniques used in machine reading comprehension [9]. Augmented Reality is a product of its time. Bilal Simsek analyzed the application of this technology to storybooks[10]. This paper provides an in-depth understanding of intelligent assessment methods.

3 Methods

3.1 Theoretical Basis

Content and examination syllabus relevance analysis is to assess whether the teaching content meets the requirements of the examination syllabus, so as to ensure the consistency between the teaching objectives and the examination standards. Reading comprehension content validity aims to assess the quality and applicability of the reading materials to ensure that they can effectively help students improve their reading comprehension skills.

3.2 Intelligent Evaluation Methods

This paper combines qualitative and quantitative analysis to conduct intelligent evaluation. In qualitative analysis, in-depth insights are obtained through expert review, interviews, and other means. Quantitative analysis presents results based on survey and experimental data. There is ambiguity in the interpretation of experts, and to construct

a model consistency matrix, it is necessary to judge the consistency of the matrix. The method is:

$$S_i = \sum_{k=1}^n x_{ik} \quad (1)$$

$$S_{ik} = \frac{S_i - S_k}{2n} + 0.5 \quad (2)$$

The mathematical expression for establishing a comprehensive weight quantification model for objective English reading comprehension questions is as follows:

$$W = \sum_{i=1}^s \sum_{k=1}^T x_{ik} v_{ik} \quad (3)$$

W is the quantified value of the comprehensive difficulty weight.

In this paper, we will use intelligent algorithms to study the intelligent assessment test system for English reading comprehension.

3.3 Assessment Modeling

The Rasch single-parameter model requires only one positional parameter, and by placing the subject and the topic on a common, continuous scale, the responses of the tested group on a given topic are inferred accordingly. The formula is calculated as:

$$\ln\left(\frac{q_{in}}{1 - q_{in}}\right) = y_i - c_m \quad (4)$$

A certain topic is m and the difficulty is c_m . The Rasch model can predict students' performance and can effectively help teachers and students to organize learning materials according to teaching objectives and interests.

The factors affecting the difficulty of English reading comprehension are mainly divided into text difficulty factor and topic difficulty factor.

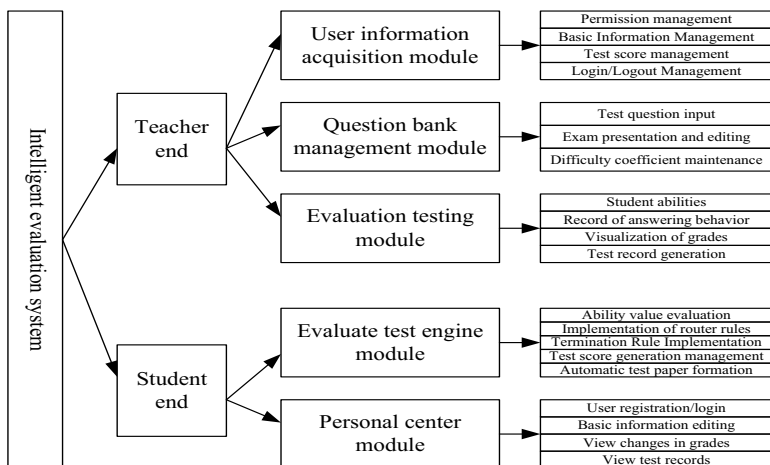


Fig. 1. Module composition of English reading comprehension assessment system

As shown in Figure 1, the intelligent assessment system designed in this paper is mainly divided into the teacher's end and the student's end, and the teaching.

3.4 Investigation Experiment

This paper randomly selected 100 students from two classes of first year non English majors, including 35 males and 65 females. In the absence of incomplete questionnaires or all questionnaire options being the same, the actual number of participants is 90% and the response rate is 90%. The question papers for this test were reading questions from previous year's CET-6 papers and the total marks were set at 250. At the end of the test, the scores of these 100 students were categorized according to their level of performance (low, medium, and high), where those in the low band scored below 150, those in the medium band scored in the range of [150,200], and those in the high band scored above 200. The comparison of English reading scores at three levels is shown in Table 1:

Table 1. Comparison of English reading scores at three levels

	Number of people	Lowest score	Maximum score	Standard deviation
Low score	23	42	144	9.3
Medium score	67	165	200	8.1
High score	10	204	237	8.9

There were more people in the middle band and fewer in the low and high bands. Their standard deviation was 9.3 at the highest and 8.1 at the lowest. There was a difference between the two scores of high, middle and low groups. Pearson correlation analysis was conducted on the subject students' CET-6 English reading scores and the overall self-assessment mean scores, and the results showed that the p-values were all

0. In addition, it was calculated that the students' self-assessment results were found to show a positive correlation (0.42-0.55) with the reading scores.

4 Results and Discussion

4.1 Experimental Preparation

In order to evaluate the accuracy and validity of the system's measurement of the subjects' English reading comprehension ability, the selection of the subjects should match the difficulty specification of the questions in the question bank. In order to achieve the purpose of application and evaluation, three measurement methods will be used in the measurement process, namely, electronic test papers, questionnaires and school English test papers.

4.2 Experimental Test

According to the intelligent assessment system, the students were divided into the experimental group and the control group. The experimental group utilized the intelligent assessment system to assess the students' reading comprehension of CET-6, while the control group used the previous paper-based test paper approach to assess the reading comprehension level. In this paper, the results of a total of 10 students (5 experimental groups and 5 control groups) were selected to be presented as the experimental results. Firstly, the function of the intelligent assessment system was tested, and secondly, the validity of the system was analyzed according to the results of the experimental control.

4.3 Analysis of Results

As shown in Figure 2, this paper finds that the accuracy results of the user information acquisition module, question bank management module, evaluation testing module, and personal center module of the system in Figure 3 (a) are all above 90%, while the coverage rate of Figure 3 (b) is all above 94%.

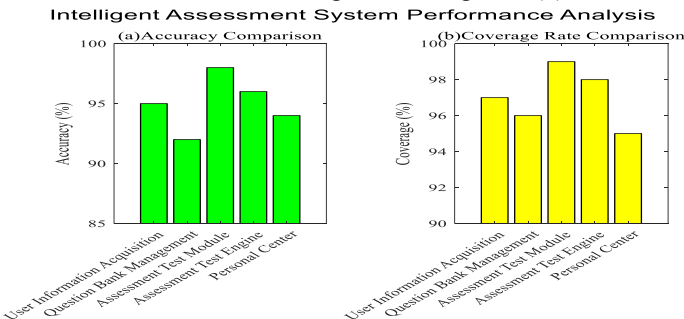


Fig. 2. Analysis of the accuracy and coverage of the intelligent assessment system

Reading Comprehension Level Improvement Comparison

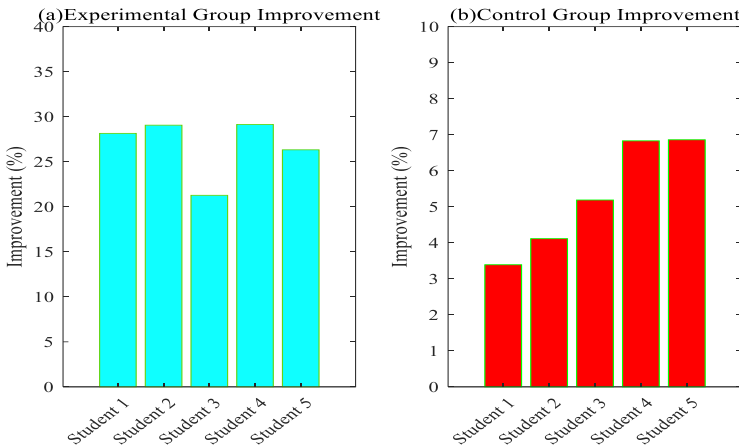


Fig. 3. Comparison of improvement in reading comprehension level

As shown in Figure 3, this paper finds that in Figure 3 (a), the reading comprehension ability of the five students in the experimental group has improved by more than 20%, indicating that the intelligent evaluation system has good effectiveness. The reading comprehension ability of the control group of 5 students in Figure 3 (b) improved by less than 8%, indicating that traditional assessment tests are still beneficial for improving student performance.

5 Conclusion

This paper designs an intelligent assessment system for English CET-6 reading comprehension, constructs a system of factors influencing the difficulty of English reading comprehension, and adopts the method of investigation and experiment to present the results of students' self-assessment. The results of the experimental test show that the reading comprehension ability of the students trained by applying the intelligent assessment system has been greatly improved compared with the traditional method. However, the data of this experiment is not rich enough to be applied to all scenarios. Therefore, this paper should also improve the robustness of the data.

References

1. Paulo Pirozelli, Marcos Menon José, Igor Cataneo Silveira, Flávio Nakasato Cação, Sara-jane M. Peres, Anarosa A. F. Brandão, Anna H. R. Costa, Fábio Gagliardi Cozman: Benchmarks for Pirá 2.0, a Reading Comprehension Dataset about the Ocean, the Brazilian Coast, and Climate Change. *Data Intell.* 6(1): 29-63 (2024).
2. Liqaa Habeb Al-Obaydi, D. R. Rahul, Marcel Pikhart: The effect of online oral reading on reading comprehension, reading anxiety, and classroom anxiety among EFL learners. *Educ. Inf. Technol.* 29(3): 2841-2855 (2024).

3. Ivan Mardini, Christian G. Quintero M., César Vilorio-Nuñez, Winston Spencer Percybrooks Bolivar, Heydy Robles, Karen Villalba: A deep-learning-based grading system (ASAG) for reading comprehension assessment by using aphorisms as open-answer-questions. *Educ. Inf. Technol.* 29(4): 4565-4590 (2024).
4. Nawras Khudhur, Aryo Pinandito, Yusuke Hayashi, Tsukasa Hirashima: Investigating the Efficacy of Partial Decomposition in Kit-Build Concept Maps for Reducing Cognitive Load and Enhancing Reading Comprehension. *IEICE Trans. Inf. Syst.* 107(5): 714-727 (2024)
5. Sophie Gruhn, Eliane Segers, Jos Keuning, Ludo Verhoeven: Understanding variation in children's reading comprehension: A dynamic approach. *J. Comput. Assist. Learn.* 40(2): 876-901 (2024).
6. Lucia Mason, Angelica Ronconi, Barbara Carretti, Sara Nardin, Christian Tarchi: Highlighting and highlighted information in text comprehension and learning from digital reading. *J. Comput. Assist. Learn.* 40(2): 637-653 (2024).
7. Anurag Shukla, Kavyansh Chourasia, Gazal Jain, Venkanna U.: DeepInsight: a CNN-based approach for machine reading comprehension in query answering systems and its applications. *Multim. Tools Appl.* 83(1): 3313-3333 (2024).
8. Pengming Wang, M. M. Kamruzzaman, Qing Chen: Machine reading comprehension model based on query reconstruction technology and deep learning. *Neural Comput. Appl.* 36(5): 2155-2170 (2024).
9. Samreen Kazi, Shakeel Ahmed Khoja, Ali Daud: A survey of deep learning techniques for machine reading comprehension. *Artif. Intell. Rev.* 56(S2): 2509-2569 (2023).
10. Bilal Simsek, Bekir Direkçi: The effects of augmented reality storybooks on student's reading comprehension. *Br. J. Educ. Technol.* 54(3): 754-772 (2023).

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

