A Study on the Content Validity of the 2025 National College Entrance Examination English Reading Comprehension

Boyu Guo

School of English Studies, Xi'an International Studies University, Xi'an, Shaanxi, China 710000

Abstract: This study analyzes the content validity of reading comprehension sections in the 2025 National College Entrance Examination (Paper I, Paper II) and Beijing Autonomous Paper. Drawing on Bachman and Palmer's test task characteristics framework and incorporating China's English Curriculum Standards for Senior High Schools (2017 Revised Edition) and official examination syllabus, the research systematically evaluates task features through dual dimensions: discourse input characteristics and expected response characteristics. Findings reveal that the reading comprehension tests demonstrate strengths in genre diversity, thematic coverage, text difficulty gradation, and reading skill assessment, aligning with standardized testing requirements. Nevertheless, certain papers exhibit opportunities for improvement in discourse selection balance and higher-order thinking skill evaluation. Based on these results, specific recommendations for test design improvements are proposed.

Keywords: English college entrance examination; reading comprehension; content validity; discourse input; expected responses

1.Research Background

The National College Entrance Examination (NCEE) is a selective examination administered to qualified high school graduates and individuals with equivalent academic ability in mainland China. The 2025 NCEE represents the first nationwide implementation of both reformed curriculum standards and a restructured English examination framework.

Since 2004, several provincial education departments have developed their own test papers based on the national syllabus, a practice that has expanded over time. This move responds to national policies encouraging local proposition, aims to mitigate the risks associated with a single national exam, accommodates regional differences, and supports quality-oriented education and reform. However, autonomous test design also introduces challenges in quality control, security, and fairness. As a result, a growing number of provinces are reverting to using national-level papers. The 2025 English NCEE employs three types of test papers: National New Curriculum Standard Paper I (National Paper I), National New Curriculum Standard Paper II (National Paper II), and Provincial Autonomous Papers. Specifically, Hebei, Jiangsu, Zhejiang, Anhui, Fujian, Jiangxi, Shandong, Henan, Hubei, Hunan, and Guangdong use National Paper I. Beijing, Tianjin, and Shanghai use Provincial Autonomous Papers. And the remaining provinces and municipalities adopt National Paper II.

Reading comprehension is a critical component for assessing language comprehension, information processing, and logical reasoning skills. Its significant weighting in the NCEE underscores its importance. Aligned with curriculum reforms and key competency frameworks, NCEE English reading assessment has shifted from evaluating discrete language knowledge toward measuring comprehensive language application skills. The English Curriculum Standards for Senior High Schools (2017 Edition, Revised) define four key competencies—language proficiency, cultural awareness, cognitive skills, and learning capacity—with the goal of integrating moral education and socialist core values into English instruction.

Content validity serves as a core metric of test quality and directly impacts the fairness and scientific rigor of the examination. Given the high stakes of the NCEE for millions of test-takers, content validity is of paramount importance. In recent years, both academic and public attention to the NCEE has increased, spurring continued research into its design.

This study analyzes the reading comprehension sections of the 2025 NCEE National Paper I, National Paper II,

and Beijing Autonomous Paper using Bachman and Palmer's task characteristics framework, with a focus on discourse input features and expected response characteristics[1]. The research evaluates their alignment with the examination syllabus and provides an empirical basis for future test development and instructional practice.

2.Theoretical Framework

Bachman and Palmer's (1996) framework of language task characteristics provides an essential theoretical foundation for language testing research[1]. Its two core dimensions—discourse input and expected response—offer a systematic approach to analyzing the content validity of reading comprehension tests. Discourse input focuses on linguistic features and material presentation, such as genre, topic, text length, and difficulty. Expected response concerns test-takers's reactions to reading tasks, particularly the coverage and distribution of reading skills. Based on this framework, numerous domestic and international scholars have conducted in-depth studies on the content validity of reading comprehension tests.

In discourse input research, attention is often directed toward the diversity of textual features and their impact on validity. Dong Manxia (2010) observed that narrative and expository texts dominated the NCEE English National Papers, with argumentative and practical texts being underrepresented[2]. Peng Yingying et al. (2017), in their analysis of the 2016 NCEE English National Papers and several Provincial Autonomous Papers, found that although autonomous papers showed greater topic variation, some exhibited limited genre diversity or uneven text length distribution[3].

Research on expected responses concentrates mainly on the distribution of reading skills. The General English Curriculum Standard for Higher Secondary Schools and the NCEE syllabus stipulate six core reading skills: understanding the main idea, comprehending specific information, inferring word meanings, making judgments and inferences, and understanding text structure and the author's viewpoint. Gu Xiangdong and Wang Qiuyan (2008), through analysis of the 2004–2006 NCEE English reading tests, found that texts during this period generally featured diverse genres and topics conforming to syllabus requirements, with well-controlled text length and vocabulary[4]. The questions primarily assessed simple inferencing skills, and text difficulty increased annually. However, concerns were noted: some tests had unevenly distributed items with over-concentrated points in certain paragraphs, undermining the holistic assessment of comprehension. Some answer options also lacked precision or deviated from the original text, potentially compromising validity. Similarly, Xiao Yunan and Lin Nan (2015) identified inadequate skill balance and coverage in the Hunan Autonomous Paper relative to syllabus requirements[5].

In summary, Bachman and Palmer's Task Characteristics Framework offers systematic theoretical and methodological support for content validity research in reading comprehension tests.

3.Research Methods

3.1 Research Objects

There are two sets of national papers in the 2025 national college entrance examination, each of which has four pieces of reading comprehension with 15 questions and a total score of 37.5. The Beijing autonomous paper has four pieces of reading comprehension with 14 questions and a total score of 28 points. Therefore, the object of this study is the 12 articles in 3 sets of 2025 English NCEE papers with a total of 44 sub-questions. There are three sets of provincial autonomous papers, namely the Beijing autonomous paper, the Shanghai autonomous paper and the Tianjin autonomous paper. In view of the fact that the Shanghai autonomous paper as well as the Tianjin autonomous paper have not been announced at present, our research object is the two sets of national papers as well as the Beijing autonomous paper for a total of three sets of papers, each set of questions with four articles in the reading comprehension section. The Beijing Paper contains 14 questions worth 28 points in total, while the other two papers (National Paper I and II) each comprise 15 questions with a total value of 37.5 points.

3.2 Data Analysis

This study takes the 2025 national college entrance examination English reading comprehension as the research object, and maintains consistency with Dong Manxia, Gu Xiangdong and Wang Qiuyan in the selection of the research object. In terms of the construction of the research framework, this study draws on the above scholars' analytical dimensions of reading comprehension tests, and develops the synchronic analysis of content validity in terms of two dimensions, namely, discourse input and expected response.

Specifically, in the discourse input dimension, this study focuses on the following four key indicators: the distribution of text genres , the diversity of topic types, the control of text length, and the level of text difficulty. In the expected response dimension, the study focuses on analyzing the examination of core reading skills, including the ability elements of comprehending the main idea, acquiring specific information, inferring word meanings, judging and reasoning, understanding the structure of the text, and grasping the author's point of view and attitude. Among them, the article length and Flesch Reading Ease of reading passages were calculated by Microsoft Office Word.

4.Results and Discussion

This study adopts a systematic analysis method to examine the reading comprehension test of the 2025 College Entrance Examination English National Paper I, National Paper II and Beijing Paper in a multidimensional way. Specifically, the study delves into the following two key dimensions. First, at the discourse input level, the study focuses on dissecting the textual features of the test. Second, at the level of expected responses, the ability test of the test questions are examined systematically. This analytical framework not only ensures the comprehensiveness of the study, but also effectively reveals the characteristics of reading tests under different propositional modes.

4.1 Discourse Input

4.1.1 Article length

As shown in Table 1, on the whole, the average length of reading comprehension articles in the three sets of English papers of the 2025 NCEE is between 300-350, and the overall length of articles is between 1230-1390, with the Beijing paper having the highest average length and total length of articles. The reason for this may be that the provinces and cities where the questions are written independently are economically and educationally more developed regions, and the overall learning ability and learning level of the students are higher, so the vocabulary assessment requirements for the candidates are a little bit higher. However, in general, they are in line with the requirement that the reading comprehension part of the test syllabus should not be less than 900 words. Liu Runqing and Han Baocheng (2004) have also mentioned that the reading comprehension material should be controlled as much as possible at about 200-500 words, and about 300 words are more desirable for tests above the intermediate level[6]. Extremely long or short texts may compromise the validity of reading ability assessment for NCEE candidates. Therefore, it can be seen that both the national papers and the Beijing paper meet this requirement and are of reasonable length.

Table 1 Statistics on the length of reading comprehension articles of the three sets of NCEE English test papers in 2025 (unit: words)

	Part A	Part B	Part C	Part D	Average length	Total Length
NP I	319	301	320	331	318	1271
NP II	275	310	284	362	308	1231
Beijing paper	234	363	402	385	346	1384

4.1.2 Article Difficulty

The difficulty of an article is divided into the calculation of Flesch Reading Ease and the amount of new words. For the calculation of readability, this paper will refer to the Flesch Reading Ease formula, which is currently

recognized and used by many scholars. The higher the score, the easier it is to understand. The Flesch readability formula is: 206.835 - (1.015 x ASL)- (84.6 x ASW). Where ASL is the average sentence length (number of words divided by the number of sentences) and ASW is the average number of syllables per word (number of stanzas divided by the number of words). The Flesch Reading Ease index categorizes reading difficulty into seven levels based on numerical scores: very difficult, difficult, fairly difficult, standard, fairly easy, easy, and very easy. These levels correspond to the following U.S. student grade levels: college level, high school level, 8th–9th grade, 7th–8th grade, 6th grade, 5th grade, and 4th grade.

The readability statistics of the reading comprehension texts of the three sets of NCEE English papers in 2025 are shown in Table 2, which shows that overall the average Flesch Reading Ease of a single text ranges from 52-64. It means that the reading difficulty of the three sets of NCEE English papers is at the fairly difficult and standard level. Among them, National Paper I has the lowest Flesch Reading Ease, indicating that National Paper I is fairly difficult for NCEE candidates. The National Paper I Flesch Reading Ease and the Beijing Paper ease of reading are equivalent to the level of American high school students. The National Paper II is equivalent to the reading level of American high school students in grades 7-8. This discrepancy arises because National Paper I (NP I) is primarily administered in the more educationally developed central and eastern regions, where test-takers generally demonstrate higher overall English proficiency. Consequently, NP I test questions are designed to align with their expected knowledge base. In contrast, National Paper II (NP II) targets the western regions and some provinces with relatively limited educational resources. To ensure relative fairness, NP II's difficulty level is accordingly adjusted to be lower. Therefore, for the college entrance examination, which is a high-stakes selective examination set up for general higher education enrollment, the difficulty is more appropriate and has a certain degree of differentiation.

In this study, operationally defined as new words are those glossed with Chinese translations in the reading comprehension sections of the NCEE English papers. As indicated in Table 2, the proportion of new words across the three NCEE papers ranges from 0.49% to 0.79%. These figures all fall well below the 3% threshold stipulated in the curriculum standards for reading comprehension vocabulary. Notably, NP I exhibits the highest proportion (0.79%), while NP II shows the lowest. This pattern aligns with the earlier finding that NP I is administered in provinces with relatively richer educational resources, thus necessitating higher-level reading texts.

Table 2 Statistics on the readability of reading comprehension articles in the three sets of NCEE English papers in 2025

	Flesch Reading Ease				Mean value	Average level of ease of reading	New words (%)
	Part A	Part B	Part C	Part D			
NP I	50	62.4	59.4	39.1	52.725	Fairly difficult	0.79%
NP II	64.2	73.3	54.7	59.9	63.025	Standard	0.49%
Beijing paper	51.8	62.7	65.2	57	59.175	Fairly difficult	0.65%

4.1.3 Article Topics

Analysis reveals that the reading comprehension passages across the three NCEE English papers cover nine thematic categories: environmental protection, education, social life, popular science, culture, health, advertisements, personal experience, and philosophy of life. The National Paper I covers four thematic areas: Environment, Education, Social Life, and Science.

The National Paper II includes four themes: Environment, Science, Culture, and Health. The Beijing Paper comprises four categories: Social Life, Advertisement, Personal Experience, and Philosophy. Among these, environmental protection, social life, and popular science emerge as the most frequent themes, while the remaining topics are evenly distributed. National Papers I and II demonstrate a stronger emphasis on environmental and popular science themes, whereas the Beijing Paper shows a clear preference for themes related to personal growth, such as

philosophy of life and personal experience.

This thematic diversity aligns with the requirements outlined in the NCEE outline, which emphasizes exposure to varied subject matter. Such breadth serves a meaningful pedagogical role: by engaging with diverse topics, candidates can broaden their cognitive perspectives, transcend disciplinary boundaries, and form cross-domain knowledge networks—effectively reducing thematic narrowness in test preparation. For instance, National Paper I integrates topics such as transportation emission reduction and synthetic fuel technology to foster interdisciplinary thinking. National Paper II incorporates real-world issues such as food waste experiments to enhance practical relevance, encouraging students to engage with authentic problems. Meanwhile, the Beijing Paper promotes affective development through discourses on life philosophy, facilitating value reflection and personal growth.

Overall, the use of four distinct themes per paper ensures broad thematic coverage across all sets.

4.1.4 Article Genre

Analysis of the 2025 NCEE English papers reveals that expository texts were the most prevalent genre in the reading comprehension sections, comprising 33.3% of the total. Narrative and argumentative texts followed, each accounting for 25%. This distribution reflects considerable genre diversity across the examination. Notably, while both National Paper II and the Beijing Paper included all four genres, that is expository, narrative, argumentative, and practical. National Paper I did not contain any practical texts. It is therefore recommended that future examinations strive for a more balanced representation of genres across all test papers.

4.2 Expected Responses

Table 3 data indicates that across 44 reading comprehension questions in the 2025 NCEE English papers, Understanding Specific Information in the Text constituted the predominant skill tested (52.3%), followed by Making Judgments and Inferences (15.9%). Understanding the Main Idea and Understanding the Author's Intention, Views, or Attitude accounted for 11.4% and 9.1% respectively. While all six predefined skills were represented across the three papers, analysis reveals that National Paper II lacked questions assessing Inferring Meanings of Words and Phrases from Context and Understanding the Author's Intention, Views, or Attitude. We therefore recommend incorporating these skill assessments in future test designs. Furthermore, National Paper II's heavy emphasis on Understanding Specific Information suggests uneven skill distribution, indicating a need for broader skill coverage.

Notably, both National Paper I and the Beijing Paper prioritized Understanding Specific Information. This aligns with Dong Manxia's (2010) assertion that this skill represents a fundamental candidate competency. Additionally, questions evaluating Making Judgments and Inferences—a higher-order cognitive skill requiring analysis of contextual clues, synthesis of textual evidence, and logical deduction—comprised a significant proportion across all papers.

Table 3 Distribution of reading skills for reading comprehension articles in the 3 sets of NCEE English papers in 2025

	NP I	NP II	Beijing Paper	Total (%)
Understanding the Main Idea	1	1	3	11.4
Understanding Specific Information in the Text	8	10	5	52.3
Inferring Meanings of Words and Phrases from Context	1	0	1	4.5
Making Judgments and Inferences	2	3	2	15.9
Understanding the Basic Structure of the Text	1	1	1	6.8
Understanding the Author's Intention, Views, or Attitude	2	0	2	9.1

5.Conclusions and Implications

Regarding discourse input features, the study indicates generally appropriate characteristics in the reading comprehension sections of the 2025 NCEE English papers. Text lengths comply with syllabus specifications. Flesch Reading Ease scores place passages between Difficult and Standard levels, corresponding to regional educational resource disparities and serving differentiation objectives. New words control satisfies curriculum standards. Thematic coverage encompasses nine topics with environmental protection, social life, and popular science each constituting 16.8 % dominance, while other topics show balanced distribution at 8.3% each. Expository texts represent the primary genre at 33.3 %, followed by narrative and argumentative texts both at 25 percent. Practical texts remain underrepresented at 16.7 % with National Paper I entirely excluding this genre.

Concerning expected responses, Understanding Specific Information in the Text predominated at 52.3 %, confirming its foundational status according to Dong(2010). Conversely, higher-order skills demonstrated notably lower representation: Understanding the Main Idea accounted for 11.4 %, Understanding the Authors Intention Views or Attitude constituted 9.1 %, while Inferring Meanings of Words and Phrases from Context and Understanding the Basic Structure of the Text received minimal coverage. National Paper I and the Beijing Paper assessed all six target skills, whereas National Paper II omitted contextual vocabulary inference and authorial attitude analysis. Furthermore, National Paper IIs emphasis on detail identification with ten questions indicated significant skill distribution imbalance.

Analysis reveals distinct propositional tendencies. National Papers prioritized socio-scientific themes including environmental protection and popular science. The Beijing Paper emphasized personal growth content such as life philosophies and individual experiences. Crucially, Making Judgments and Inferences received substantial attention across all papers, representing a higher-order skill demanding contextual analysis and evidentiary synthesis.

Primary improvement areas include achieving balanced topic distribution, expanding practical texts coverage, optimizing skill allocation through increased higher-order focus, and addressing National Paper IIs skill gaps and imbalances. Study limitations involve sample representativity constraints due to unavailable Tianjin and Shanghai papers, necessitating broader jurisdictional coverage in future research.

References:

[1] Bachman,L.F.,&Palmer,A.S.(1996).*Language Testing in Practice*.Oxford University Press.

[2] Dong,M.X.(2010).Diachronic analysis of reading comprehension tests in the NCEE English National Papers. *Foreign Language Teaching in Schools(Secondary Section)*,33(02),31–37.

[3] Peng,Y.Y.,Gu,X.D.,&Huang,J.(2017).Evaluating the quality of 2016 NCEE English reading comprehension tests:A task characteristic framework analysis based on Bachman and Palmer.*Educational Measurement and Evaluation*,(03),57–64.

[4] Gu,X.D.,&Wang,Q.Y.(2008).Content validity analysis of reading comprehension tests in NCEE National Papers and provincial autonomous papers.*Testing Research*,(03),102–114.

[5] Xiao,Y.N.,&Lin,C.(2015).Content validity of Hunan autonomous papers' reading comprehension tests(2010 – 2014).*Educational Measurement and Evaluation(Theoretical Edition)*,(03),44–49+28.

[6] Liu,R.Q.,&Han,B.C.(2004).*Language Testing and Its Methods*.Foreign Language Teaching and Research Press.