

---

# Cambridge A2 Key for Schools Assessment: A Comprehensive Analysis of Basic English Proficiency Assessment

Shuang Liu\*

Taiyuan Richen Academy, Taiyuan 030032, Shanxi, China

*\*Author to whom correspondence should be addressed.*

**Copyright:** © 2025 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), permitting distribution and reproduction in any medium, provided the original work is cited.

---

**Abstract:** This study conducted an in-depth analysis of the Cambridge A2 Key School English Test (KET). Through the analysis and sorting of the literature and KET materials and the evaluation of construct validity, design rationality, and scoring reliability, it was concluded that the KET test structure is consistent with the CEFR A2 level test. KET mainly assesses learners' basic communication skills, and the test materials emphasize that the test questions should be close to the authenticity of real-life tasks and meet international standards. The analysis of this paper reveals that KET emphasizes communication and practical language application, thus reflecting its strong practicality. Through the research and analysis, it is found that its limitations are mainly insufficient coverage in writing, a lack of cognitive adaptability for young learners, and an assessment of cross-cultural communication. Therefore, this study recommends the introduction of a diagnostic system assisted by computerized adaptive testing (CAT) to improve the accuracy of assessment and develop a wider and more accurate assessment scope. The results of this study will provide valuable insights for the assessment of learners' abilities in basic education assessment at this stage and in the future. The emphasis on strong authenticity, standardization, and consideration of the needs of young learners will become the focus of related research on this topic.

**Keywords:** KET; English language assessment; CEFR; Communicative competence; Basic education

---

**Online publication:** March 26, 2025

## 1. Introduction

Cambridge A2 Key for Schools (formerly known as KET for Schools) is a foundational English language qualification test designed specifically for school-age learners<sup>[1]</sup>. It is an ideal starting point for young learners and provides them with the confidence to pursue higher-level qualifications such as B1 Preliminary and B2 First. The exam is aligned with Level A2 of the Common European Framework of Reference (CEFR 2001), indicating that candidates can use English to communicate effectively in simple situations<sup>[2]</sup>.

According to the Cambridge Assessment English website ([www.cambridgeenglish.org](http://www.cambridgeenglish.org)), the A2 Key for Schools exam consists of three main components:

(1) Reading and writing (60 minutes): This section includes seven parts with a total of 32 questions. It tests students' ability to understand simple written information such as signs, brochures, and short articles, as well as the ability to write simple sentences and short paragraphs.

(2) Listening (30 minutes): Candidates listen to a variety of spoken materials, such as announcements and short conversations, and answer questions to demonstrate their understanding of specific information and the overall meaning of the texts. The listening section includes five parts with 25 questions. The listening materials are presented at a reasonably slow pace to accommodate the basic level of the exam.

(3) Speaking (8–10 minutes): The speaking test is conducted in pairs or groups of three. Students are assessed on their ability to interact with the examiner and their partner, including comparing, describing, and expressing opinions. The assessment criteria include grammar and vocabulary, pronunciation, and interactive communication skills.

## **1.1. Research background and significance**

### **1.1.1. The importance of English assessment in basic education**

English, as the global language of communication, holds a significant position in basic education. Effective English assessment is crucial as it provides valuable feedback for teaching and learning. Tosuncuoglu stated that assessments can provide both teachers and students with information about the level of knowledge <sup>[3]</sup>, skills, difficulties in foreign language learning and which activities and methods are the most useful. It is also a means of evaluation of student activities and can be used for a clear criterion or scale. The opinion of Bachman and Damböck is that language assessment is as a process of using the results of an assessment to arrive at interpretations about students' language ability <sup>[4]</sup>, and to make decisions in order to help bring about beneficial consequences for students, teachers, the school, and perhaps other individuals and institutions. Furthermore, China's English curriculum standards aim to promote learning and teaching through assessment, and to integrate assessment into the entire process of subject teaching <sup>[5]</sup>.

### **1.1.2. The typicality of KET in language assessment**

The A2 Key for Schools exam (KET) is the entry-level examination within the Cambridge English Qualifications framework and is internationally recognized. Its design and implementation serve as a model for other language assessments. The extensive application of the KET in basic education makes it an important case study for exploring the theories and practices of language assessment. This research on the KET can contribute to the advancement of the field of language assessment.

## **1.2. Research questions and methodology**

### **1.2.1. Core research questions: Construct validity, design rationality, and scoring reliability**

This study focuses on the construct validity of the KET exam, exploring whether it accurately measures the intended

language abilities. It also examines the rationality of the exam design, assessing whether the task settings and question types are appropriate for the candidates' actual proficiency levels. Furthermore, the study investigates the scientific reliability of the scoring process, ensuring that the scoring is fair and accurate.

### **1.2.2. Research approach based on literature analysis**

This research employs a systematic review of domestic and international literature on the KET exam to summarize previous research findings as well as related studies and identify gaps in the existing studies. By conducting a comprehensive examination of the official KET materials, this study provides an in-depth analysis of the exam's theoretical constructs, assessment design, and scoring system. The findings are used to propose research conclusions and recommendations for improvement.

## **2. Theoretical framework and construct analysis**

### **2.1. Positioning of language proficiency within the CEFR framework**

#### **2.1.1. Mapping analysis of A2-level language proficiency descriptors**

The A2 level of the Common European Framework of Reference (CEFR) describes the language proficiency of basic users in everyday life situations. The listening, speaking, reading, and writing tasks of the KET exam are highly aligned with the A2-level proficiency descriptors, ensuring that the exam content is consistent with international standards. Namely, in the speaking section, the A2 level requires candidates to engage in simple everyday communication. The KET exam assesses candidates' communicative abilities by simulating real conversational scenarios, thereby reflecting the A2-level proficiency requirements of the CEFR.

#### **2.1.2. Application of the communicative language ability model**

The communicative language ability (CLA) model proposed by Bachman and Palmer emphasizes the integration of language knowledge and language use abilities, suggesting that effective language proficiency should encompass both the structural knowledge of the language (such as vocabulary and grammar) and the ability to use this knowledge in communicative contexts [6]. The KET exam, in its design, fully incorporates this principle by assessing both language knowledge and language use abilities, thereby reflecting the application of the CLA model. This integration is fully reflected in the reading module of the KET exam. The main purpose of the reading module is to assess candidates' language knowledge through various types of texts, using vocabulary and grammatical structures that can ensure they can recognize and understand. At the same time, their reading comprehension skills are assessed through task design, which includes the ability to interpret texts and derive meaning from them. Language knowledge and the ability to use the language can be effectively combined by this dual approach, which is consistent with the CLA model's emphasis on communicative competence.

### **2.2. Construct features of the KET exam**

#### **2.2.1. Construct definition oriented towards everyday communicative competence**

Aligned with the CEFR, the Cambridge A2 assessment is designed to evaluate learners' basic communicative

competence in real-life contexts. The content of the exam is closely related to real-life situations, such as shopping, traveling, and making friends, ensuring that candidates can apply their acquired language knowledge in practical contexts. For example, the question-and-answer format is mainly reflected in the oral mode. The purpose is to simulate interpersonal communication and test the candidates' daily communication skills, which reflects the construction-oriented and communication-oriented nature of the examination.

### **2.2.2. Embodiment of task-based language testing (TBLT) theory**

Task-based language testing (TBLT) theory emphasizes the assessment of language ability through authentic tasks. Long and Doughty pointed out that the most important of the communicative approaches is task-based teaching [7]. The way of structuring learning is to let learners complete meaningful tasks, and to a certain extent, emphasize language use and communicative ability rather than grammatical accuracy. The KET exam incorporates this theory by designing authentic tasks in the reading, writing, and speaking sections. Specifically, the reading comprehension tasks include advertisements and announcements, the writing tasks involve writing notes and emails, and the speaking tasks feature conversational exchanges. These tasks not only assess candidates' language knowledge but also their ability to use language in real-life situations, thereby enhancing the practicality and effectiveness of the exam.

## **2.3. Controversial aspects of construct representation**

### **2.3.1. Balance between language knowledge and communicative competence**

While the Cambridge A2/KET emphasizes practical communication [8], Weir cautioned that such exams may still overemphasize discrete language knowledge at the expense of communicative competence [9]. Canagarajah pointed out that the ability of job applicants to "travel" between different language communities with different language usage norms should be assessed [10]. To achieve this goal, a new paradigm is called for, which emphasizes language awareness, sociolinguistic skills, and negotiation skills. For instance, the KET's speaking tasks, though structured around social interactions, often rely on scripted prompts that may not fully assess a learner's ability to negotiate meaning or adapt to unexpected conversational turns. Similarly, the separation of reading, writing, listening, and speaking into discrete sections may fail to capture the integrated nature of real-world language use.

### **2.3.2. Absence of intercultural communicative competence**

While the Cambridge A2/KET effectively assesses basic communicative competence in predictable, everyday contexts, it operates largely in the absence of intercultural communicative competence (ICC). For instance, the oral tasks are useful but only in scripted monocultural interactions. The learners' ability to negotiate meaning or adapt to different cultural norms cannot be assessed [11]. Similarly, the absence of reflective tasks in the KET exam suggests that there is no assessment of critical cultural awareness, which, according to Byram, is a key component of ICC [12]. This limitation undermines the test's relevance in today's globalized world, where English is increasingly used as a

lingua franca among speakers of diverse linguistic and cultural backgrounds.

### **3. Analysis of examination design features**

#### **3.1. Overall structural design**

##### **3.1.1. Module composition and weight allocation (reading and writing/ listening/ speaking)**

The KET exam is composed of three modules: Reading and Writing, Listening, and Speaking, with respective weightings of 50%, 25%, and 25%. This weight allocation reflects the significance of reading and writing in foundational English proficiency while also accommodating the assessment of listening and speaking skills. The Reading and Writing module evaluates candidates' language knowledge and written expression abilities, the Listening module assesses their listening comprehension skills, and the Speaking module examines their oral communicative abilities. Together, these modules form a comprehensive English proficiency assessment system <sup>[13]</sup>.

##### **3.1.2. Appropriateness of test duration and task density**

The KET exam has a total test duration of approximately 1 hour and 40 minutes, with a moderate task density. The Reading and Writing section includes multiple subtasks, the Listening section comprises various listening materials, and the Speaking section involves several conversational segments, ensuring that candidates have sufficient time to complete the examination tasks. Each reading and writing part has a clear time limit. Candidates must complete the task within the specified time, and there is no situation where there are too many tasks and they cannot be completed.

#### **3.2. Analysis of task types and their characteristics**

##### **3.2.1. Multimodal input design in the reading module**

The reading module incorporates a multimodal input design that can be informed by Mayer's multimedia learning principles <sup>[14]</sup>, which emphasize the benefits of combining text with relevant visuals to enhance comprehension. This design enriches the presentation of reading materials and assesses candidates' ability to comprehend different types of information. For instance, advertisements and announcements often contain materials such as charts and pictures in the reading section. These materials test the ability to use text and graphic information to understand the content, thereby improving the comprehensiveness and practicality of the reading task.

##### **3.2.2. Authenticity of interactive assessment in the speaking test**

Authenticity in language testing refers to the degree to which test tasks resemble real-life language use. According to Bachman and Palmer, authenticity is a key quality of effective language assessments and is closely tied to interactivity, which involves the engagement of the test-taker's language ability, background knowledge, and cognitive strategies <sup>[15]</sup>.

The speaking test in the KET exam evaluates candidates' oral communicative ability through interaction with an

examiner and another candidate. This form of interaction is highly authentic, simulating real-life communication scenarios and assessing candidates' ability to engage in conversation and respond to unexpected situations. For example, the authenticity of the oral test's interactivity is mainly reflected in the fact that candidates communicate with examiners and other candidates, express their own opinions, and answer questions from others. This authenticity mainly occurs in the Q&A and group discussion sessions of the oral test.

### **3.3. Task design principles**

#### **3.3.1. Simulation of real-life situations**

The task design of the KET exam places a strong emphasis on simulating real-life situations, enabling candidates to apply their acquired language knowledge in practical contexts. The exam content encompasses everyday scenarios such as shopping, traveling, and making friends, thereby enhancing the authenticity of the tasks. For example, the note-taking task in the writing section is designed to simulate real situations. In the task, candidates will use the language knowledge they have learned to effectively express their intentions. Therefore, candidates must write notes based on the actual situation.

#### **3.3.2. Gradient control strategy for cognitive load**

The KET exam task design takes into account the gradient control of cognitive load, progressively increasing task difficulty from simple to complex. This strategy helps candidates gradually acclimate to the exam's difficulty level, thereby enhancing the fairness and effectiveness of the assessment. For instance, the matching of simple information to complex reading comprehension reflects the increasing difficulty of the reading module. Therefore, candidates can complete the tasks of the simple module first and then complete the complex tasks to enhance their reading ability.

#### **3.3.3. Implementation of cultural neutrality**

The KET exam also adheres to the principle of cultural neutrality in its task design, avoiding cultural biases that might affect candidates. The exam content includes scenarios from diverse cultural backgrounds, ensuring that candidates can complete tasks regardless of their cultural context. For example, the cultural customs of different countries are mainly the topics of the oral test. Respecting the other party's cultural background and reflecting the principle of cultural neutrality are what candidates should keep in mind when expressing their opinions.

## **4. Scoring system and validation of validity**

### **4.1. Scoring criteria system**

#### **4.1.1. Integration of analytic and holistic scoring**

The scoring criteria system for the KET exam integrates both analytic and holistic scoring. Analytic scoring evaluates

the specific performance of each task, while holistic scoring provides a comprehensive assessment of the examinee's overall performance. For example, in oral assessment, examiners conduct analytic scoring based on the examinee's pronunciation, grammar, vocabulary, fluency, and other aspects, while also providing a holistic evaluation of the examinee's overall oral performance to ensure the comprehensiveness and accuracy of the scoring.

#### **4.1.2. Empirical validity research on the oral scoring scale**

Bachman and Palmer pointed out that the definition of construct validity can be described as “the extent to which the indicators of the ability or structure we want to measure can be interpreted through the interpretation of specific test scores.” Therefore, they believe that the foundation of language testing is construct validity, as the purpose of the test and how to reasonably assess the language ability of test takers based on the test scores are determined by it. Bachman and Palmer further proposed the Assessment Use Argument (AUA) model [16]. They emphasized that this model underscores the systematic and structured nature of validity argumentation. The difference between the AUA model and traditional validity theory is that this model will take the test development process as the core part of validity argumentation, and take the expected results of the test as the starting point of validity argumentation. In short, test developers need to clarify the expected results of the test, which should be preset in the test design stage, and finally construct the test content and scoring criteria around the obtained test results.

The oral scoring scale has undergone empirical validity research to ensure the scientific nature and reliability of the scoring criteria. The research involves collecting a large amount of oral examination data to analyze the reliability and validity of the scoring scale, thereby continuously optimizing the scoring standards. For instance, this indicates that the rating scale has good reliability and validity. An important indicator of whether it can accurately assess the oral ability of candidates is whether the rating scale has a high degree of consistency in the ratings of different examiners.

### **4.2. Reliability assurance mechanisms**

#### **4.2.1. Dual calibration system for examiner training**

The KET exam has established a dual calibration system for examiner training to ensure consistency in scoring. Examiners are required to pass rigorous assessments during the training process to ensure their familiarity with the scoring criteria and procedures. For instance, to ensure the accuracy of the scoring, learners will be graded on a large number of sample exercises and asked to compare their scores with the correct answers to calibrate their own scores.

#### **4.2.2. Complementarity of machine scoring and human scoring**

The KET exam integrates both machine scoring and human scoring in the evaluation process, fully leveraging their complementary strengths of both methods. Machine scoring is used for objective questions, while human scoring is applied to subjective questions. For example, the objective questions in the reading module are scored by machines to

ensure objectivity and accuracy, whereas the oral component is scored by professional examiners to ensure flexibility and fairness.

### **4.3. Validity verification research**

#### **4.3.1. Validation framework based on Messick's unified view of validity**

The validity verification of the KET exam employs a validation framework based on Messick's unified view of validity, which assesses the validity of the examination from multiple dimensions. The validation includes content validity, construct validity, and criterion-related validity, among others. For instance, in order to ensure the content validity of the exam, requirements may be made for expert review and practical exams to verify the consistency of the exam content with the CEFR A2 level requirements.

#### **4.3.2. Research on washback effects**

Alderson and Wall pioneered the study of washback in language testing, challenging the assumption that tests inherently influence teaching and learning <sup>[17]</sup>. The washback effects of the KET exam have been validated through empirical research. Gu and Saville pointed out that the positive washback effect of KET exams has been generated in teaching in China <sup>[18]</sup>. Their research emphasized that KET exams are generally considered a motivational tool. Students can participate in English learning more actively through this tool, which provides them with motivation to learn to a certain extent. In other words, the researchers found that KET will pay more attention to the practicality of communication skills. The goal of English education in China also emphasizes the practicality of English communication skills. This practicality can cultivate learners' self-confidence and improve learners' classroom participation.

## **5. Comprehensive evaluation and suggestions for improvement**

### **5.1. Summary of strengths**

#### **5.1.1. Authenticity and communicative orientation of tasks**

The task design of the KET exam emphasizes authenticity and communicative orientation, effectively assessing candidates' ability to use language in real-life contexts. The test content is close to daily life, and the task design is close to reality. Using the learned language knowledge to communicate is a requirement for candidates and the main purpose of the test <sup>[19]</sup>.

#### **5.1.2. Standardization and international recognition**

The KET exam is characterized by a high degree of standardization and international recognition. To ensure the fairness of the exam, the exam design must be scientific, and the scoring standards must be strict and clear. This will gain international recognition and provide candidates with globally recognized proof of English proficiency.



## **5.2. Analysis of existing issues**

### **5.2.1. Limited genre coverage in the writing module**

The writing module of the KET exam has limitations in terms of genre coverage. Simple writing styles such as notes and emails are the main scope of writing tasks. Due to the lack of assessment of other forms, the overall evaluation level of candidates' writing ability is also limited.

### **5.2.2. Insufficient adaptation to the cognitive characteristics of young learners**

The KET exam falls short in adapting to the cognitive characteristics of young learners. The main reason that can lead to difficulties during the exam is that the exam content and task design are too high-level for young test takers. For instance, certain tasks in the reading module require a level of comprehension that may exceed the cognitive abilities of young learners. This discrepancy can hinder their performance and affect the fairness of the examination [20].

## **5.3. Recommendations for optimization**

### **5.3.1. Introduction of computerized adaptive testing (CAT) technology**

The introduction of computerized adaptive testing (CAT) technology is recommended to dynamically adjust the difficulty of the test based on candidates' ability levels, thereby enhancing the precision and efficiency of the examination. For example, CAT can adaptively adjust the difficulty of subsequent questions based on candidates' performance on previous items, ensuring that candidates are tested at an appropriate level of difficulty. This approach not only improves the fairness of the examination but also enhances its validity.

### **5.3.2. Development of a multidimensional diagnostic feedback system**

A multidimensional diagnostic feedback system is proposed to provide candidates with detailed and comprehensive feedback on their performance. This system would offer feedback across multiple dimensions, including language knowledge, language use, and communicative competence, helping candidates identify their strengths and areas for improvement. For instance, the diagnostic feedback system could generate detailed score reports that highlight candidates' performance in each dimension and provide targeted recommendations for improvement, thereby assisting candidates in enhancing their English proficiency.

### **5.3.3. Enhancement of intercultural communicative elements**

It is recommended to enhance the integration of intercultural communicative elements within the test by incorporating more intercultural communicative contexts. This would allow for the assessment of candidates' intercultural communicative competence. For example, the speaking section could include topics related to

intercultural communication, requiring candidates to demonstrate their understanding and respect for different cultures during the interaction. This enhancement would strengthen the examination's focus on intercultural communicative ability.

## **6. Conclusion**

### **6.1. Summary of key findings**

This study has conducted an in-depth analysis of the construct, design, and scoring system of the KET exam. It was found that the KET exam has significant strengths in terms of task authenticity, communicative orientation, standardization, and international recognition. However, it also has limitations, such as the restricted genre coverage in the writing module and insufficient adaptation to the cognitive characteristics of young learners <sup>[21]</sup>.

### **6.2. Implications for educational assessment reform in basic education**

The research on the KET exam offers valuable insights for the reform of English language assessment in basic education in our country. It is recommended that future assessments focus on the authenticity and communicative orientation of tasks, enhance the standardization and international recognition of assessments, and pay attention to the cognitive characteristics of young learners. Additionally, the assessment should incorporate more elements of intercultural communicative competence.

### **6.3. Future research directions**

Future research could further explore the application effects of the KET exam in different educational contexts and investigate how to better integrate the KET exam with teaching practices to promote students' English proficiency. Moreover, research could focus on the application of the KET exam in computerized adaptive testing and multidimensional diagnostic feedback systems, thereby advancing the development of language assessment technologies.

## **Disclosure statement**

The author declares no conflict of interest.

## **References**

- [1] A2 Key Exam Format, 2025, Cambridge English, viewed March 10, 2025, <https://www.cambridgeenglish.org/exams-and-tests/key/exam-format/?skill=vocabulary>
- [2] Council of Europe, Council for Cultural Co-operation, Education Committee, Modern Languages Division, 2001, Common European Framework of Reference for Languages: Learning, Teaching, Assessment, Cambridge University

Press, Cambridge.

- [3] Tosuncuoglu I, 2018, Importance of Assessment in ELT. *Journal of Education and Training Studies*, 6(9): 163–167.
- [4] Bachman L, Dambock B, 2018, *Language Assessment for Classroom Teachers*, Oxford University Press, Oxford.
- [5] Chinese Ministry of Education, 2022, *English Curriculum Standards for Compulsory Education (2022 Edition)*, Beijing Normal University Publishing Group, Beijing.
- [6] Bachman LF, Palmer AS, 1996, *Language Testing in Practice: Designing and Developing Useful Language Tests (Vol. 1)*, Oxford University Press, Oxford.
- [7] Long MH, Doughty C, (Eds.), 2009, *The Handbook of Language Teaching*, Wiley-Blackwell, Chichester, 373.
- [8] Cambridge Assessment English, 2023, *KET Handbook for Teachers*, Cambridge English, Cambridge.
- [9] Weir CJ, 2005, *Language Testing and Validation*, Palgrave MacMillan, Hampshire, 10.
- [10] Canagarajah S, 2006, Changing Communicative Needs, Revised Assessment Objectives: Testing English as an International Language. *Language Assessment Quarterly: An International Journal*, 3(3): 229–242.
- [11] Canagarajah AS, 2006, Negotiating the Local in English as a Lingua Franca. *Annual Review of Applied Linguistics*, 26: 197–218.
- [12] Byram M, 1997, *Teaching and Assessing Intercultural Communicative Competence*, Multilingual Matters, UK.
- [13] Gu X, Saville N, 2016, Twenty Years of Cambridge English Examinations in China: Investigating Impact from the Test-Takers' Perspectives, in *Assessing Chinese Learners of English: Language Constructs, Consequences and Conundrums*, Palgrave Macmillan UK, London, 287–310.
- [14] Mayer RE, (Ed.), 2005, *The Cambridge Handbook of Multimedia Learning*, Cambridge University Press, Cambridge.
- [15] Bachman LF, Palmer AS, 2010. *Language Assessment in Practice: Developing Language Assessments and Justifying Their Use in the Real World*, Oxford University Press, Oxford.
- [16] Bachman LF, 1990, *Fundamental Considerations in Language Testing (Vol. 87)*, Oxford University Press, Oxford.
- [17] Alderson JC, Wall D, 1993, Does Washback Exist? *Applied Linguistics*, 14(2): 115–129.
- [18] Gu Y, 2012, *English Curriculum and Assessment for Basic Education in China*, in *Perspectives on Teaching and Learning English Literacy in China*, Springer, Dordrecht, 35–50.
- [19] Harding L, 2014, Communicative Language Testing: Current Issues and Future Research. *Language Assessment Quarterly*, 11(2): 186–197.
- [20] Lazaraton A, 2002, *A Qualitative Approach to the Validation of Oral Language Tests (Vol. 14)*, Cambridge University Press, Cambridge.
- [21] Weir CJ, Vidakovic I, Galaczi ED, 2013, *Measured Constructs: A History of Cambridge English Examinations, 1913–2012 (Vol. 37)*, Cambridge University Press, Cambridge.

**Publisher's note**

Whioce Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.