

Cambridge

English Skills Test

General

Cambridge English Skills Test General Overview



CAMBRIDGE
English

This overview outlines the overall purpose of Cambridge English Skills Test *General* and *Business*, and the Cambridge approach to test validity, demonstrating how we ensure Cambridge English Skills Test is fit and valid for the purposes it was designed.

Test purpose

Cambridge English Skills Test is a modular online multi-level test of English language proficiency that reports results on an internationally recognised reference framework called the Common European Framework of Reference for Languages (CEFR).

The *General* version can be used for admissions, progress measurement, or as an exit test for higher education institutions and pathway programs. It can also be used for recruitment for roles in a non-business-specific environment where the employee needs strong English skills to perform their role but will not need specialist business terminology.

The *Business* version is designed for students entering the workplace at graduate level, or employees using the test for professional development or promotion. The aim of this version of the test is to assess whether a candidate's language ability is at a level where they are ready to engage in professional activities and operate successfully in workplace environments.

Both test versions are designed to have a positive influence on language teaching and learning. The following key features of Cambridge English Skills Test provide a sound foundation for the test results to be used in the contexts mentioned above:

- During **test construction**, tasks for all skills are carefully designed and based on CEFR descriptors and research into cognitive and contextual factors.
- Rigorous **test content review** performed by language testing professionals ensures high-quality test questions which assess critical language knowledge and skills of stakeholders' interest.
- **Trial studies** to confirm that the test assesses language knowledge and skills that are essential for real-world communication. For instance, approximately 4,000 participants of a trial study on the English Language Skills Test reported that the test assesses essential real-world knowledge and skills.
- **Regular checks** on test reliability of our computer-adaptive Reading and Listening modules. These checks show that both modules yield highly reliable test scores, with reliability estimates of 0.94 for Reading and 0.93 for Listening.
- **Quality assurance and monitoring** of examiners and the automarker for Writing and Speaking components to ensure results are fair and unbiased.
- The Writing component of the Cambridge English Skills test is marked by an **auto-marker** developed based on machine learning technology with the Institute of Automated Language Teaching and Assessment (ALTA) at the University of Cambridge. We work with researchers from the Department of Engineering, Computer Laboratory and the Department of Theoretical and Applied Linguistics to create the next generation of AI-empowered

assessment and learning tools for English language learners. Research has indicated that the automated scores are highly correlated with those awarded by certified human examiners, which are considered the gold standard.

- The Speaking component of the Cambridge English Skills test adopts a **hybrid approach to marking** in which candidates' oral responses are marked by either an auto-marker or human examiners. When the automarker indicates low confidence in marking, the test response is flagged up and escalated to human examiners. Our research has found that hybrid marking resulted in a good agreement with exclusive human marking with 95.6% exact agreement and 100% adjacent agreement on the CEFR.
- Standard setting studies are conducted to ensure **alignment of the tests to the CEFR**. The most recent studies confirmed the proper alignment between Reading and Listening scores and the CEFR framework as intended. The Writing and Speaking scores have also been reviewed by panels of experts for establishing alignment. The positive findings from these studies empower score users of the Cambridge English Skills Tests to interpret the results with confidence and make informed decisions based on CEFR.

The Cambridge English Skills Test validity framework

Frameworks allow for the systematic description of test tasks for the purpose of language test design, test validation and research. The theoretical framework selected to guide the test evaluation process for Cambridge English Skills Test is Weir's (2005) *socio-cognitive framework* for language test validation¹. The framework allows for a systematic and comprehensive critical evaluation of construct definition and operationalisation using an evidence-based approach.

The framework is described as *socio-cognitive* in that "the abilities to be tested are demonstrated by the mental processing of the learner (the cognitive dimension); equally, the use of language in performing tasks is viewed as a *social* rather than a purely linguistic phenomenon" (Taylor, 2011, p.25). The framework is very much in line with the use of language for social purposes as defined in the CEFR. It defines language ability as a combination of both cognitive abilities and the interaction of these abilities with the surrounding social and contextual factors.

The framework consists of several components that together present a unified approach to collecting validity evidence for a given test (Figure 1).

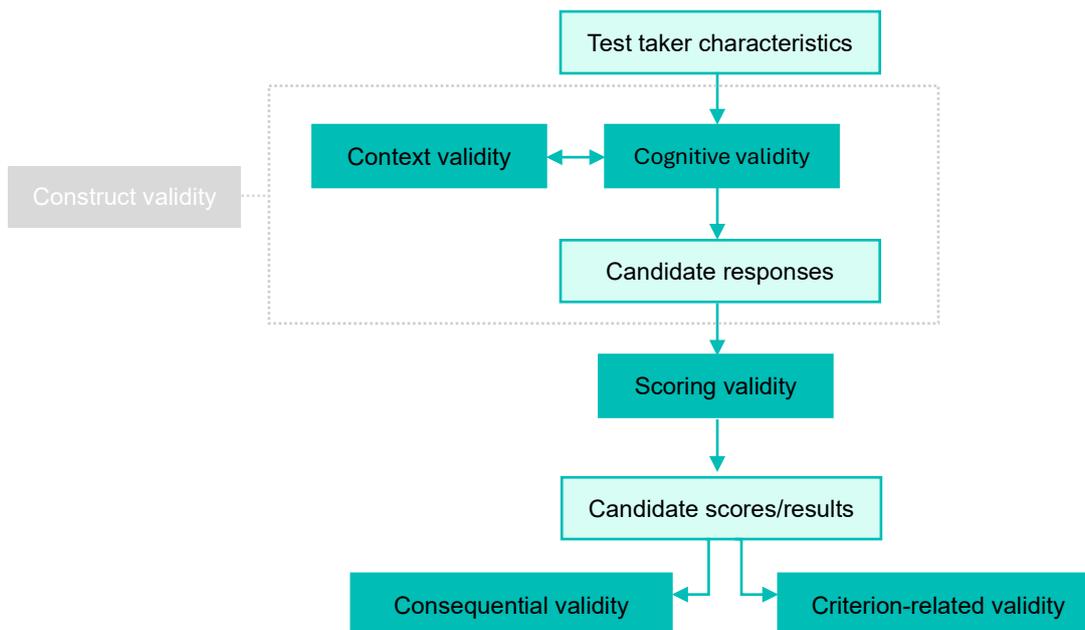


Figure 1 A socio-cognitive framework (based on Weir 2005, pp. 44-47)

In our evaluation process, we focus on five main components of validity as presented in the socio-cognitive framework, namely cognitive, context, scoring, criterion-related and consequential validityⁱⁱ. These guide our investigation of test validity and are defined as follows:

- **Cognitive validity:** it examines the type of cognitive processing that is activated by the test task(s) and the extent to which the cognitive processes required to complete the tasks are appropriate for the language level of the learners and reflective of real-world mental processes.
- **Context validity:** traditionally used to refer to the content coverage of tasks, the term is used more broadly to refer to characteristics of test tasks and the extent to which they are appropriate and fair to learners taking the tests.
- **Scoring validity:** in the Weir (2005) framework, scoring validity is used as a superordinate term to include all aspects of reliability and refers to the extent to which we can depend on the scores which result from a given test.
- **Criterion-related validity:** it can relate to three basic types of comparisons (Khalifa and Weir, 2009, pp.7-8): a connection between test scores and external criteria which purports to measure the same ability, demonstrative evidence of qualitative and quantitative equivalence across different test versions or linkage to an external standard (e.g. the CEFR).
- **Consequential validity:** it looks into the uses of the test results, and whether these are valid ones. It investigates the impact of the test at different micro and macro levels: on learners, on the classroom, on the school and on educational systems, and on society as a whole.

ⁱ This approach has proven not only to be theoretically sound but also practically useful, evidenced in a number of validation activities and research projects undertaken by Cambridge for over two decades and documented in a series of ‘construct’ volumes on writing (Shaw & Weir 2007), reading (Khalifa & Weir, 2009), speaking (Taylor, 2011) and listening (Geranpayeh & Taylor, 2013)

ⁱⁱ It should be noted that construct validity is not listed explicitly in the model, but is represented by the box surrounding the interplay between context validity (the content of the test) cognitive validity (what learners need to complete the tasks on the test) and *candidate responses* (how learners perform on the test).

Further reading

Geranpayeh, A. & Taylor, L. B. (Eds.). (2013). *Examining listening: Research and practice in assessing second language listening*. Studies in Language Testing 35. Cambridge, UK: Cambridge University Press.

Khlaifa, H. & Weir, C. J. (Eds.). (2009). *Examining reading: Research and practice in assessing second language reading*. Studies in Language Testing 29. Cambridge, UK: Cambridge University Press.

Shaw, S. D. & Weir, C. J. (Eds.). (2007). *Examining writing: Research and practice in assessing second language writing*. Studies in Language Testing 26. Cambridge, UK: Cambridge University Press.

Taylor, L. (Ed.). (2011). *Examining speaking: Research and practice in assessing second language speaking*. Studies in Language Testing 30. Cambridge, UK: Cambridge University Press.

Weir, C. J. (2005). *Language Testing and Validation*. Basingstoke: Palgrave Macmillan.

Flexible, reliable assessment for the decisions you need to make

Cambridge
English
Skills Test

Discover more:
cambridge.org/skillstest



Find out more at
cambridge.org/english

We believe that English can unlock a lifetime of experiences and, together with teachers and our partners, we help people to learn and confidently prove their skills to the world.

Where your world grows