

Critical Evaluation of C1 Advanced English Test Reading Based on Socio-Cognitive Framework

Qianhui Zhu

Guangzhou College of Technology and Business, Guangzhou, Guangdong 510800

Abstract: Due to the rising mobility of the population, it is imperative to introduce appropriate tools for assessing non-native speakers' English-language skills and predict their academic or professional performance in order to obtain further progression. Due to this significance, this paper aims to investigate the construct validity of reading section in the C1 Advanced English Test devised by the Cambridge English Language Assessment. For validation, Weir's socio-cognitive framework is applied to examine the reading tasks of the 2021 sample paper. The finding shows that construct of the reading tasks are valid and scientific, thus suitable for organisations and candidates to use it as a reference of their language proficiency.

Keywords: CAE Reading; Construct validity; Socio-cognitive framework

1. Introduction

C1 Advanced English Test (CAE) is a proficiency test launched by Cambridge English Language Assessments in 1991. In January 2015, the CAE test has been revised, merging the Reading and Use of English sections into a single paper. The aim of 2015 revision of CAE is to ensure its suitability for higher education studies and personal professional development (Zeronis & Geranpayeh, 2015). The purpose of the paper is to investigate construct validity of the reading session to see whether it fits into the purpose the exam expects and meets the need of the test takers.

2. Critical evaluation on CAE reading

2.1 Rationale

For the purpose of a more comprehensive and systematic analysis of the test, Weir's socio-cognitive framework is adopted. In this framework, construct validity seems act as an umbrella term because context validity, cognitive validity, scoring validity, consequential validity and criteria validity are subordinated (Weir, 2004). This paper is going to focus on the test itself, thus consequential and criterion-related validity which is looking outside of the world of the test will not be discussed in details.

2.2 Context validity

In terms of the context validity (or content validity), it refers to the extent to which the test content is a representative of the skills meant to be measured (Hughes, 2003). It concerns with the purpose, response format, order of items, time constraints, text length and so on (Weir, 2004).

The CAE Test is designed to assess individual English proficiency for candidates whose first language is not English, providing an opportunity for them to demonstrate language competence in workplace, academic courses and life in an English-speaking country. Dogalus (2000) states that the content of the test should authentically reflect the tasks in real situation, which helps to make inferences about the language ability of the test takers. The texts selected in the CAE reading section, such as magazine, journals, non-literary books, leaflets, brochures and so on, are text types that test takers commonly encounter in work and study.

As to the response format, various types are involved, including Multiple Choice, Cross-Text Multiple Matching, Gapped Text, Multiple Matching Questions. The combination of different response formats is beneficial to elicit different levels of cognitive processing.

2.3 Cognitive Validity

As to the cognitive validity, it is about the extent to which the test items arouse the variety and depth of the cognitive process in a reading test which aims to mimic the real-world academic process (Weir & Bax, 2012). In order to give in-depth analysis of cognitive process in different tasks, a cognitive processing model of reading suggested by Khalifa and Weir's (2009) is employed. Here, examples extracted from the sample paper will be used to illustrate the mental process that the reading activates.

Example 1

You are going to read four reviews of a book about how architecture can affect the emotions. For questions 37 – 40, choose from the reviews A – D. The reviews may be chosen more than once.

Which reviewer

Has a different opinion from the others on the confidence with which de Botton discusses architecture?

37. (University of Cambridge Local Examinations Syndicate, 2021)

The above Cross-text multiple matching is designed to compare the opinions and attitudes across four short texts, and match the texts to four appropriate statements. Candidates are not only required to do local comprehension at the word, syntax and sentence level, but also comprehend beyond the sentences to get the gist of the four short texts. As Weir and Bax (2012) state that careful reading is a general term which involves processing at the sentence, intersentential, text, and multi-text levels and it is an important ability enabling students to be capable of college level reading comprehension.

Example 2

You are going to read an extract from a magazine article. Six paragraphs have been removed from the extract. Choose from the paragraphs A – G the one which fits each gap (41 – 46). There is one extra paragraph which you do not need to use.

Scottish Wildcat

On my living-room wall I have a painting of a wildcat by John Holmes of which I am extremely fond. It depicts a snarling, sitting animal, teeth bared and back arched: a taut coiled spring ready to unleash some unknown fury.

41. (University of Cambridge Local Examinations Syndicate, 2021)

However, the physical differences are tangible. The wildcat is a much larger animal, weighing in some cases up to seven kilos, the same as a typical male fox. The coat pattern is superficially similar to a domestic tabby cat but it is all stripes and no spots. The tail is thicker and blunter, with three to five black rings. The animal has an altogether heavier look.....

The response format used in this part is Gapped-Text Task in which candidates are supposed to read the text and paragraphs in order to determine which paragraph is the most appropriate for each other. In this task, a higher level of cognitive process is required as the candidates need to recognise the macro- or micro- propositions, as well as how they are related (Khalifa and Weir, 2009). In this sense, candidates, firstly, decode the meaning of words and group them into phrases then sentences to build up the micro-proposition. Then, they assemble the information accessed in the previous stages to build up the macro-proposition. Therefore, careful global and local readings are activated in tackling this type of question.

Example 3

You are going to read a magazine article in which five career consultants give advice about starting a career. For questions 47–56, choose from the consultants (A–E). The consultants may be chosen more than once.

Which consultant makes the following statements?

47. Keep your final objective in mind when you are planning to change jobs.

(University of Cambridge Local Examinations Syndicate, 2021)

The above is multiple-matching task, which intends to test the ability to identify specific information, detail, opinion and attitude. The process model activated in this task is careful local reading and search reading. In this part, nearly all the answers are contained in single sentences. Candidates, guided by the predetermined topic, establish sentence level propositional meaning (Khalifa and Weir, 2009). Initially, candidates decode the meaning of “final objective” and “change job” in the statement, and search them in the text. Accordingly, synonyms like “goal” and “move on from your first job” can be found in the corresponding sentence — “But in the early stages of your career you need a definite strategy for reaching your goal, so think about that carefully before deciding to move on from your first job.” By accessing the lexical meaning and the syntactic meaning, candidates can establish the meaning at the sentence level to choose the correct answer.

2.4 Scoring Validity

Reliability is regarded as one form of validity evidence and also termed as scoring validity. It concerns about the degree to which the test produces consistent and stable results (Weir, 2004).

All the tasks in CAE reading are objectively scored items, eliciting an appropriate answers from the choices. This contributes to the overall reliability of the test, as internal consistency and maker reliability can be guaranteed. In addition, the ESOL has used a number of approaches and statistical measurements to monitor the stability of the test. For instance, ESOL has established a Local Item Banking System to store test items and analyze their difficulties based on an underlying scale (Weir, 2004). In addition, Cronbach's Alpha and Standard Error of Measurement (SEM) are also used to check the internal consistency of the test. For the former, the value of 0.8 and above is regularly achieved (Cambridge English, 2016), which can be regarded acceptable for a component. The latter, it reflects a candidate's score may differ if they were to take the test on another occasion, and the lower the value the more reliable the test is. According to the research carried by Elliott, Lim, Galaczi and Calver (2012), the SEM value of trial reading test was around 3.0.

Conclusion:

To sum up, by looking into the context validity, cognitive validity and scoring validity of the 2015 revision of the CAE reading component, it is evident that the construct of the test is valid and scientific, ensuring it is one of the suitable exams that candidate can take to demonstrate their reading competence in higher education and career development.

References:

- [1]Cambridge English (2016) Principles of Good Practice: Research and Innovation in
- [2]Language Learning and Assessment. Cambridge: Cambridge English Language Assessment.
- [3]Douglas, D. (2000). Assessing languages for specific purposes. Cambridge University Press.
- [4]Elliott, M, Lim, G, Galaczi, E and Calver, L. (2012). FCE and CAE Construct Validation Study (Part 2), Cambridge: Cambridge English internal report.
- [5]Hughes, A. (2003). Testing for Language Teachers. Cambridge: Cambridge University Press.
- [6]Khalifa, H., & Weir, C. J. (2009). Examining Reading: Research and Practice in Assessing Second Language Reading (Studies in Language Testing) (1st ed.). Cambridge University Press.
- [7]University of Cambridge Local Examinations Syndicate. (2021). CAE Sample Paper 1. Cambridge: Cambridge English Language Assessment.
- [8]Weir, C. (2004). Language testing and validation: An Evidence-Based approach (research and practice in applied linguistics) (2005th ed.). Palgrave Macmillan.
- [9]Weir, C., & Bax, S. (2012). Investigating Learners' Cognitive Processes During a Computer-Based CAE Reading Test. Research Notes, 47, 3–13.
- [10]Zeronis,R., & Geranpayeh, A. (2015). Continuity and innovation: Updating FCE and CAE. Research Notes, 62, 3–5.