valid instrument for a broad sample of potential examinees. If used at all, it should be interpreted with great caution and considered suggestive at best with regard to an individual's entrepreneurial potential.

**REVIEWER'S REFERENCE**

[68]

**Examination for the Certificate of Competency in English.**

**Purpose:** Designed as a test of general language proficiency in a variety of contexts; it assesses linguistic, discoursal, sociolinguistic, and pragmatic elements of the English language.

**Population:** Nonnative speakers of English.

**Publication Dates:** 1995–2011.

**Acronym:** ECCE.

**Scores, 5:** Writing, Listening Comprehension, GVR (Grammar/Vocabulary/Reading), Speaking, Overall Score.

**Administration:** Group.

**Forms, 4:** A, B, C, D.

**Price Data:** Available from publisher.

**Time:** (30) minutes for Listening; (80) minutes for Grammar/Vocabulary/Reading; (30) minutes for Writing; (10-15) minutes for Speaking.

**Comments:** Designed for nonnative speakers of English who would like official documentary evidence of intermediate proficiency in the English language, particularly for academic and professional purposes.

**Author:** The University of Michigan.

**Publisher:** Cambridge–Michigan Language Assessments.

**Review of the Examination for the Certificate of Competency in English by ANTONY JOHN KUNNAN, Professor, California State University, Los Angeles, CA, and ELVIRA WAGNER, Assistant Professor, Temple University, Philadelphia, PA:**

**DESCRIPTION.** The Examination for the Certificate of Competency in English (ECCE) is a partially group-administered paper-and-pencil standardized assessment of general language proficiency at the high-intermediate level of English as a foreign language. Scores are used as evidence of high-intermediate competence in English for personal, public, education, and occupational purposes. The test is administered to tens of thousands of test takers around the world at Cambridge Michigan Language Assessment (CaMLA) centers; CaMLA is a not-for-profit collaboration between the University of Michigan and the University of Cambridge. The majority of test takers whose data are included in the November–December 2006 test report were from Europe—mainly from Greece—with smaller numbers from South America, Latin America, and Asia. The ECCE has four sections: listening; grammar, vocabulary, and reading (GVR); writing; and speaking. In total, the ECCE takes 2 hours and 35 minutes, although because of the need to schedule individual speaking administrations, test takers usually are required to be at the examination site for longer.

The listening section has 50 multiple-choice questions with three response options each and lasts for 30 minutes. Test takers record their responses on a bubble sheet. There are two parts to the listening section. Part 1 contains 30 items that involve listening to short recorded conversations followed by questions. For Part 2 of the listening section, a recorded radio interview is played. The interview is broken into segments, and groups of questions follow each segment.

The GVR section has 100 multiple-choice items with four response options each and takes 80 minutes. Part 1 is the grammar component with 35 items in which an incomplete sentence is followed by a choice of words or phrases to complete it. Part 2 is the vocabulary component, again with 35 questions. Part 3 is reading, with 30 questions presented in three parts. Part 1 involves a short reading passage followed by six comprehension questions, Part 2 has short texts that are presented as advertisements and accompanied by 11 or 12 questions, and Part 3 has a longer reading passage with 12 or 13 questions.

The writing section of the ECCE takes 30 minutes during which test takers write one sample. Test takers are given a very short article to read, and they can then choose to write either a letter or an essay in response to the article. These written samples are scored using a rating scale with four categories: content and development, organization and connection of ideas, linguistic range and control, and communicative effect. The holistic rating scale uses 5 points: A (high pass), B (pass), C (low pass), D (borderline fail), and E (fail).

The speaking section of the ECCE takes approximately 15 minutes. The test taker meets individually with an examiner. There are six components to the speaking test: opening/warm-up stage 1 (conveying non-sensitive personal information), stage 2 (asking questions to get information), stage 3 (expressing an opinion and supporting it), stage 4 (elaborating on the topic), and a conclusion. The
examiner who is both the interlocutor and the rater, rates speaking in three categories: language control—grammar, language control—vocabulary, and delivery and intelligibility. The 5-point scale used is the same as for the writing section.

The listening, GVR, and writing score sheets are sent to CaMLA for scoring. The listening and GVR sections are machine-scored; the writing section is scored by CaMLA trained and certified raters.

DEVELOPMENT: The ECCE is designed as a criterion-referenced competence examination to assess whether test takers have achieved the high-intermediate English language ability in alignment with the Common European Framework of References for Languages (CEFR). The CEFR describes language ability on a 6-level scale: A1 (lower basic), A2 (upper basic), B1 (lower intermediate), B2 (upper intermediate), C1 (lower proficient), C2 (upper proficient). The ECCE assesses whether test takers meet the threshold of B2.

The ECCE was designed to assess test takers' communicative use of English. A new form of the ECCE is written for each test administration. Items are written according to predetermined format and difficulty levels and are piloted on samples representative of the intended ECCE population of test takers.

TECHNICAL.

Standardization. Because the ECCE scores are aligned with the CEFR, the test is criterion-referenced, and thus, a standardization/norm sample does not exist. The CEFR describes language ability in terms of what a language user at a particular level can do. Because the ECCE seeks to assess whether test takers meet the B2 level of the CEFR, the items are designed to assess the abilities described at the B2 level and must possess the desired psychometric properties (e.g., discrimination indices, target difficulty range, and interitem correlations). Linking items are inserted into new operational tests and are used to equate new forms with previously administered forms.

Reliability. The May 2010 ECCE Test Administration Report indicates that 64.1% of the test takers passed the May 2010 examination. For the May 2010 administration, the reliability and standard error of measurement (SEM) estimates for the listening section were .89 and .34, respectively, and the reliability and SEM estimates for the GVR section were .93 and .27. The May 2010 test report does not indicate how these reliability estimates were computed, but the November-December 2006 report stated that item response theory with a three-parameter logistic model was used to estimate the reliability coefficient.

For the writing section of the ECCE, each response is rated separately by two raters trained and certified by CaMLA. If the two raters do not reach exact agreement on the score using the 5-point holistic rating scale, the writing response is rated by a third rater. For the writing section, rater agreement figures and pass/fail agreement figures are presented in the May 2010 test report. For the May 2010 administration, the rater agreement was 75.9%, and the pass/fail agreement was 89.6%.

As described earlier, the speaking samples are scored by the test interlocutor. These raters are trained using the ECCE Oral Examiner Training Manual and an accompanying DVD with videos of test takers taking the speaking test and numerous calibration samples and benchmarks. However, the performance of speaking test examiners is monitored locally, rather than by CaMLA, and thus no reports of scoring reliability are given for the speaking section. The May 2010 test report states that recordings of speaking tests are sent to CaMLA for review of how the speaking tests are conducted by the local examiners and for the accuracy of the ratings, although there is no mention of how CaMLA reviews these recordings, nor is there information about what the standards for accuracy and reliability are. In addition, the fact that the rater is also acting as interlocutor presents possible threats to reliable scoring, as does the fact that the speaking samples are scored by just one rater.

Validity. Validity evidence has been provided to support inferences made from the ECCE. Validation studies include Liao (2007), who investigated the construct validity of the ECCE by investigating the extent to which performance on the GVR section predicts performance on the listening section. Another validation study conducted by Lu (2005) examined the differences in discourse used by non-native and native speaker examiners. The results of this study influenced subsequent oral examiner training on the ECCE. The May 2006 test report describes a study supporting the internal structure of the test. Factor analyses were conducted on the listening and GVR sections, which presented evidence in support of the models used to operationalize test items.

Fairness. An additional area of investigation that complements reliability and validity is fairness:
whether the ECCE is fair to test takers from different groups such as gender, age, and disability. DIF analyses were conducted, and about 14% of the items were flagged for DIF, with roughly half of these items favoring males and half favoring females. However, because test takers from different age groups take this test, it would be necessary to know whether the test tasks are of comparable difficulty for the different age groups of test takers. Finally, as more test takers with disabilities are taking such tests, it would be necessary to know what accommodations are provided to test takers with various disabilities and whether any studies have been conducted to ensure the accommodations provided are not biased in favor of or against test takers with disabilities. These studies would enhance the value of the ECCE.

COMMENTARY. The ECCE is a test of general language proficiency at the high-intermediate level of English as a foreign language. However, there are a number of issues that need to be considered in evaluating the ECCE’s overall usefulness. First, the ECCE is linked with the CEFR as it examines whether test takers meet the threshold of B2 in the CEFR. However, the ECCE documents provided for review do not offer a study that verifies that the test is really at the B2 level. Second, although a human interlocutor in the speaking test is valuable as it enhances the authenticity of the interaction in the test and the validity of the inferences made from such an interaction, there is a possible loss of reliability and standardization by having a human interlocutor. Therefore, a study that provides evidence that having a human interlocutor who is also the rater does not result in a loss of reliability and standardization is essential. Finally, the ECCE needs to consider accommodations for test takers with disabilities.

SUMMARY. The ECCE is a partially group-administered paper-and-pencil standardized assessment of general language proficiency at the high-intermediate level of English as a foreign language at the B2 level of the CEFR. The test is well designed and serves the purpose of assessing test takers’ communicative language competence in English.

REVIEWERS’ REFERENCES