

Language Testing

<http://ltj.sagepub.com>

Test review: College English Test (CET) in China

Ying Zheng and Liying Cheng
Language Testing 2008; 25; 408
DOI: 10.1177/0265532208092433

The online version of this article can be found at:
<http://ltj.sagepub.com>

Published by:

 SAGE Publications

<http://www.sagepublications.com>

Additional services and information for *Language Testing* can be found at:

Email Alerts: <http://ltj.sagepub.com/cgi/alerts>

Subscriptions: <http://ltj.sagepub.com/subscriptions>

Reprints: <http://www.sagepub.com/journalsReprints.nav>

Permissions: <http://www.sagepub.com/journalsPermissions.nav>

Citations (this article cites 9 articles hosted on the
SAGE Journals Online and HighWire Press platforms):
<http://ltj.sagepub.com/cgi/content/refs/25/3/408>

different and much more academically oriented than at the three lower levels.

In summary, the GEPT appears to be a solid, traditional test battery with a focus on high practicality at the three lower levels, whereas the two upper levels resemble more recent generations of academically and communicatively oriented tests. The test is widely available throughout Taiwan (at least at the lower levels) and less costly for test takers than other options like TOEFL or IELTS. Construct validation research, investigations of social and educational impact, and empirical anchoring of can-do statements appear to be the most pressing items for GEPT research.

References

- McNamara, T. F. & Roever, C.** (2006). *Language testing: The social dimension*. Oxford: Basil Blackwell.
- Pan, Y. C.** (2007). Consequences and evaluation of test use: The washback of exit requirements on tertiary English education in Taiwan. Manuscript in preparation.
- Weir, C. J. & Wu, J. R. W.** (2006). Establishing test form and individual task comparability: A case study of a semi-direct speaking test. *Language Testing*, 23(2), 167–197.
- Wu, R. & Chin, J.** (2006). An impact study of the intermediate level GEPT. Retrieved January 17, 2008, from http://www.ltcc.ntu.edu.tw/2007_annual_report/p59-p74.pdf.

College English Test (CET) in China

Ying Zheng and **Queen's University, Canada**

Liying Cheng **Queen's University, Canada**

DOI: 10.1177/0265532208092433

Test purpose: The purpose of the CET is to examine the English proficiency of undergraduate students in China and ensure that Chinese undergraduates reach the required English levels specified in the National College English Teaching Syllabuses (NCETS) (Syllabus for College English Test, 2006, p. 1). In accordance with the

Address for correspondence: Ying Zheng, Faculty of Education, Queen's University, Kingston, Ontario K7P 3N6, Canada; e-mail: 2yz13@queensu.ca

courses on the syllabus, the CET consists of three tests: Band 4 (CET-4), Band 6 (CET-6), and the CET – Spoken English Test (CET-SET).

Length and administration: The test is administered by the National College English Testing Committee on behalf of the Higher Education Department, Ministry of Education, the People's Republic of China (the P. R. China). The CET-4 and the CET-6 are held twice a year at the end of each semester. Each test takes 125 minutes to complete. The CET-SET is also held twice a year, a month before the end of each semester. The test lasts approximately 20 minutes.

Scores: Four subscores are reported in the CET report cards in addition to the total score gained out of 710 points: listening comprehension (249 points, 35%), reading comprehension (249 points, 35%), cloze or error correction (70 points, 10%), and writing and translation (142 points, 20%). The CET-SET results are reported in a graded scale, A, B, C, and D, with A representing the highest level and D representing the lowest level.

Author/publisher and Contact information: The National College English Testing Committee. The Administration Office of the National College English Testing Committee, Shanghai Mailbox 30–14, Rm. 2203, Haoran Hi-Tech Building, Shanghai Jiaotong University, Shanghai 200030, the P. R. China, Tel: +86–21–52583311, Fax: +86–21–62932756. CET website: www.cet.edu.cn.

Price: The prices for the CET vary in different provinces of China due to the local cost of living. The average is around 25 RMB for the CET-4 and the CET-6 and 50 RMB for the CET-SET (at the time of writing, the exchange rate of US dollar vs. RMB was roughly 1:7).

I General description

The CET test-takers are undergraduates in China, who are majoring in any discipline except English. Students who have completed the College English Courses Band 1 to 4, usually second-year students, take the CET-4. Those students who have completed the College English Courses Band 5 to 6 and have passed the CET-4, usually third-year students, take the CET-6. Test-takers of the CET-SET must be undergraduates who have either passed the CET-4 with a score of at least 550 or passed the CET-6 with a score of at least 520. The number of the CET test-takers is very large. In 1987, 100,000 students took part in the first CET-4 administration; in 2006 the number

soared to 13,000,000. This number makes it the largest English as a foreign language test in the world and one of the language tests that has attracted most public attention in China (Jin, 2005).

The CET has gone through three important phases during its 20-year history. The first phase was from 1987 to 1997, when fundamental components of the test were laid out. The second phase (1997–2005) featured test format changes, including the addition of compound dictation in listening and translation from English to Chinese in reading in 1997, and the introduction of the CET-SET in 1999. Further reforms in test content, formats and scoring system were piloted in 2005 and this set off a third phase of the test. The new CET was introduced across the country in 2006.

In this current phase, changes were made in the following areas: 1) the percentage of marks for listening increased from 20% to 35%; 2) the vocabulary and structure section, which used to constitute 15% of the total score, was removed; 3) new test content, such as skimming and scanning and translation, were introduced; 4) new test formats, such as banked cloze, true or false, and short answer questions, were introduced; 5) an online marking system of subjective items (writing and constructed response items) was introduced; 6) the scoring system was changed, and the concept of pass or fail was eliminated. The new scoring system has a total of 710 ($M = 500$, $SD = 70$), as compared to 100 in the old system ($M = 72$, $SD = 12$) (National College English Testing Committee, 2006).

Listening

The listening section assesses students' abilities to understand oral conversations or passages based on standard American English or standard British English (Syllabus for College English Test, 2006, p. 3). The speed of the conversation is approximately 130 words per minute (wpm) for the CET-4 and 150 wpm for the CET-6. The listening comprehension contributes 35% to the total score, within which 15% is for *listening conversations*, including short conversations (one round of conversation and a question) and long conversations (5–8 rounds of conversational exchanges and 3–4 questions), and 20% is for three *listening passages* and one *compound dictation passage*. The section is 35 minutes long.

Reading

The reading section assesses students' ability to understand English in written formats. The reading comprehension contributes

35% to the total score, within which 25% is for *Reading in Depth*, including three short passages (300–350 words per passage), and 10% is for *Skimming and Scanning*, which is made up of 1–2 long articles (approximately 1,000 words for the CET-4 and 1,200 words for the CET-6). Three item formats are employed in *Reading in Depth*: multiple choice questions, banked cloze, and short answer questions. In the banked cloze format, students are asked to select one word for each blank (10 blanks in total) from a list of 15 choices given in a word bank following the passage (Syllabus for College English Test, 2006, p. 3–4). In the short answer question format, students are asked to complete a sentence or answer questions with less than 10 words based on their understanding of the passage. *Reading in Depth* takes 25 minutes. For *Skimming and Scanning*, students have 15 minutes to go over the passage quickly (the required speed is 100 wpm for the CET-4 and 120 wpm for the CET-6), and obtain the main idea of the passage and locate important information. The item formats employed in *Skimming and Scanning* include multiple choice questions, true or false, and sentence completion.

Cloze

This section takes 15 minutes and it constitutes 10% of the total score. Students are asked to choose correct answers from four choices for the 20 blanks in a short passage (220–250 words). Students' abilities to understand and infer meanings from vocabulary, sentence, and paragraph levels are tested in this section. For the CET-6, an alternative format for this section is error correction, which asks students to identify and correct 10 mistakes embedded in a passage of about 220–250 words.

Writing and translation

This section assesses students' ability to communicate ideas using written English. It takes 35 minutes and it constitutes 20% of the total score. For writing, students are asked to write a composition of no less than 120 words for the CET-4 and 150 words for the CET-6 based on the information provided (e.g., title of the topic, outline, situation, pictures, or graphs etc). The time limit is 30 minutes. For translation, students are asked to complete five sentences by translating content given in Chinese into English in 5 minutes.

Speaking

The speaking test, an optional component for students, is composed of three sections. Section one is an interaction between 3 or 4 test-takers and 2 authorized CET-SET examiners (one interlocutor/assessor and the other assessor) for approximately 5 minutes. Section two is a 1.5-minute presentation for each individual and a 4.5-minute group discussion (approximately 10 minutes in total). Section three involves more questions from the interlocutor to further test the test-takers' oral English ability (approximately 5 minutes).

Students' performance in the speaking section is evaluated based on six criteria: 1) *Accuracy* in pronunciation, stress/intonation and use of grammar and vocabulary; 2) *Range* of vocabulary and grammatical structures; 3) *Size* (percentage) of contribution (to group discussion) made by each individual candidate; 4) *Discourse management* represented by the ability to produce extended and coherent discourse; 5) *Flexibility* in dealing with different situations and topics; 6) *Appropriateness* in the use of linguistic resources according to context (National College English Testing Committee, 2006, p. 1).

Score interpretation

This test is described as a "criterion-related norm-referenced test" (Syllabus for College English Test, 2006, p. 9). The test criteria are based on the National College English Syllabus (before 2004) and the National College English Curriculum Requirements (after 2004). It is norm-referenced because the CET scores are equated against a pre-determined norm (Yang and Jin, 2001).

II Appraisal of the CET

The following appraisal of the CET is based on Kunnan's (2004) concept of test fairness, using the five main qualities of validity, absence of bias, access, administration, and social consequences. Kunnan (2004) views fairness in terms of "the whole system of a testing practice, not just the test itself" (p. 37), which means that not only the steps in the test development process should be examined, but also the test use by multiple stakeholders and intended and unintended test consequences should be evaluated.

Validity

In the test fairness framework (Kunnan, 2004), the concept of test validity includes content representativeness, construct or theory-based validity evidence, criterion-related validity evidence and reliability. The main focus of examining test validity is to check if test items, tasks, or topics are representative of the target content and underlying traits, if test scores can be explained in accordance with external criteria, and if the inter-rater and internal consistency of marking has reached its desirable stability. The first and primary validation study of the CET was conducted by Yang and Weir (1998). The results of this study provided a comprehensive overview of the validity evidence for each section of the CET in its first phase and triggered changes to the test format and the introduction of the CET-SET.

Yang (2000) further reported that to achieve content validity in reading, the CET committee established detailed test specifications and test item construction processes. Different content areas, such as specific language skills and knowledge, are constructed and randomly selected to avoid content overlap or bias. An analysis of 300 reading comprehension items in previous test papers indicated that these CET reading items fully covered the content in the test specification. To examine the criterion-related validity of the CET-4, 50 university classroom teachers' rankings of 1,186 students' English levels were compared to their CET-4 test scores. The correlation coefficient reached 0.7. Further, Yang (2000) reported that the discriminant validity index was 0.82, which indicated that the CET-4 test scores were a strong discriminator of whether or not test-takers' English abilities reached the required abilities, thus serving the function of its norm referencing well. Validity evidence regarding the CET reading is also reported in an introspective verbal protocol study (Jin and Wu, 1998). The results demonstrated that the skills students reported in completing the CET reading are in line with the skills expected by the test designers.

It should be noted that CET validation studies have mostly been conducted for the reading section and the limited number of studies were mostly conducted around 10 years ago referring to the old version of the test. At the moment, with the introduction of the new version in 2006, there is a serious need for more research to validate the new CET, especially sections other than reading.

The internal consistency reliability coefficient for the objective items of the old version of the CET was consistently over 0.9 (Yang and Weir, 1998). A series of measures have also been taken to ensure

the inter-rater and intra-rater reliability of the writing assessors. Range-finders² and different levels of descriptors are determined beforehand by the National College English Testing Committee. Quality control of the rating process is conducted by random spot checking and constant supervision at three levels of the marking process (raters, supervisors, and directors) (Jin, in press).

It seems that one feature of this test that distinguishes itself from other English language tests elsewhere in the world is the test designers' efforts to achieve efficiency. This feature clearly bears specific Chinese characteristics: to be able to cater for the needs of an enormously large number of test-takers while maintaining acceptable levels of reliability and validity. As a major contributor to the CET design, Yang (2000) argued that without efficiency, large scale tests like the CET cannot be implemented in China. The online marking system has greatly helped the CET to achieve this goal. However, whether the efficiency of the marking system has been developed at the expense of its quality, it is yet to be determined through further studies.

Absence of bias

An examination of absence of bias should include detecting offensive content or language, unfair penalisation, disparate impact and standard setting (Kunnan, 2004). In the case of the CET, major efforts to avoid test item bias are reflected in the consideration of its contents with respect to candidates majoring in different undergraduate disciplines. As the CET test-takers are all Chinese undergraduates, cultural or ethnic bias is not expected to be a major issue.

Access

The concept of test access includes educational, financial, geographical, personal access and equipment and condition familiarity (Kunnan, 2004). The CET-4 is accessible to students who have completed the College English Courses Band 1 to Band 4 and the CET-6 to those who have finished Band 5 to Band 6. Before the new version of the CET was adopted in 2006, most universities only allowed second-year students to take the test in order to guarantee higher pass rates. With the recent test reform, Chinese universities have now adopted a loose

²The term "range-finder" refers to the prototype responses chosen to represent students' writing at specified score ranges.

practice of linking the CET performance to students' undergraduate degrees. Students can take the CET whenever they feel ready.

Administration

The CET administration follows a typical top-down structure to ensure uniformity and security. The national Higher Education Department oversees the general guidelines of the test. The implementation of the test is conducted by the National Education Examinations Authority, which supervises the administrative practices of the provincial/municipal education examinations authority. The CET administration office is responsible for technical testing-related practices.

Social consequences

An appraisal of the CET should never underestimate its social consequences due to its enormously large number of test-takers and the tradition of using examinations for selection purposes in China. The CET measures the English proficiency of undergraduate students in accordance with the College English Teaching Syllabus – thus exerting tremendous impact on university/college teaching and learning of English in China and affecting the huge numbers of stakeholders (see Cheng, 2008 for a full review).

Some Chinese researchers point out the positive social consequences of the CET as it has promoted the role of English teaching and learning at the tertiary level in China over its 20-year history (e.g., Li, 2002, Gu, 2005). At the same time, other researchers (e.g., Han, Dai & Yang., 2004; Gu and Liu, 2005) challenge the CET by pointing out that the test does not assess communicative competence as the teaching syllabus requires. Before the recent reform of the CET, the large proportion of multiple choice questions made this test an efficient one, but possible sacrifices in accurately measuring English language ability remained a concern (Zhang, 2003). Since the CET certificates had been one of the graduation requirements of undergraduates in the majority of Chinese universities for almost 20 years, test-oriented teaching was another severe problem (Gu and Liu, 2005).

This test is now turning to a direction that attempts to promote better interaction between English language teaching, learning, and testing in China. Remedies for its possible negative social influences are being put in place. For example, the introduction of the CET-SET is a strong supplement to the test battery, and it has played an active

and positive role in promoting students' communicative competence (Jin, 2000). The changes made in the CET test content, format, and scoring systems are evidence of the endeavors to bring about positive washback of the test on English teaching and learning in China. It is the expectation of the CET committee that the test can reflect and catch up with the needs of rapid economic reform and the new open-door policy (Jin, 2006; in press). Especially, it is expected that the shifts in the weightings of test components and test contents can increase the emphasis on communicative functions. It is also expected that the new scoring system can lead to a more rational utilization of the test results. However, whether or not positive washback on the teaching of English at the tertiary level in China has been accomplished clearly merits further research and discussion. Changes to the test, like everything else in China, are inevitable, and hopefully these can bring benefits to the test stakeholders, especially the test-takers.

III Summary

In a vast country like China, where the function of testing as a fair indicator of students' academic success is generally accepted (Cheng, 2008, Cheng & Curtis, in press), it is imperative to evaluate and enhance the quality of a test. The CET is a large-scale language test with distinctive Chinese characteristics. However, with the fast increasing number of Chinese students studying in English-medium universities around the world outside China, the impact of the CET on Chinese students' English learning has already gone beyond its intended use and consequences on Chinese society alone. Its impact on test-takers' learning of English needs further research. In addition, creating a test that serves the dual function of criterion-related and norm referenced testing has also been challenging for the test designers and problematic for test users. It is not an exaggeration to say that the CET is the most debated test among academics in China. In the past two decades, the CET has served as a good benchmark around which English language teaching and learning in China has revolved. A concern of the testing committee, given the numbers of the test-takers, is whether those numbers of test-takers are "a blessing or a worry" (Jin, 2005, p.30). The CET was made possible due to the rapid expansion of tertiary education in Chinese society; it, in turn, contributes greatly to the route of development in English language teaching and learning for Chinese society.

IV References

- Cheng, L. (2008). The key to success: English language testing in China. *Language Testing* 25, 15–37.
- Cheng, L. & Curits, A. (Eds.). In press. *English Language Assessment and the Chinese Learner*. New York: Routledge: Taylor's & Francis Group.
- Gu, X. (2005). Positive or negative? An empirical study of CET washback on college English teaching and learning in China. *ILTA Online Newsletter* 2 Retrieved September 5, 2007, from <http://www.iltaonline.com/newletter/02-2005oct/>
- Gu, W. & Liu, J. (2005). Test Analysis of College Students Communicative Competence in English. *Asian EFL Journal* 7, 118–33.
- Han, B., Dai, M. & Yang, L. (2004). Problems with College English Test as emerged from a survey. *Foreign Languages and Their Teaching* 179, 17–23.
- Jin, Y. (2000). The washback effects of College English Test-Spoken English Test on teaching. *Foreign Language World* 118, 56–61.
- Jin, Y. (2005). The national College English Test of China. Paper presented at the International Association of Applied Linguistics (AILA), Madison, WI.
- Jin, Y. (2006). On the improvement of test validity and test washback: The CET washback study. *Foreign Language World* 6, 65–73.
- Jin, Y. In press: The national College English Testing Committee of China. In L. Cheng & A. Curtis (Eds.), *English Language Assessment and the Chinese Learner*. New York: Routledge: Taylor's & Francis Group.
- Jin, Y. & Wu, J. (1998). Examining the testing validity of CET reading comprehension by introspection. *Foreign Language World* 2, 47–52.
- Kunnan, A. (2004). Test fairness. In M. Milanovic & C. Weir (Eds.), *Selected papers from the ALTE conference, Barcelona* (pp.27–48). Cambridge, UK: Cambridge.
- Li, J. (2002). The current College English Test in China: Problems and thoughts. *Foreign Language Education* 23(5), 33–38.
- National College English Testing Committee. (2006). *College English Test-Spoken English Test (CET-SET) Syllabus and Sample Test*. Shanghai, China: Shanghai Language Education Press.
- National College English Testing Committee. (2006). *CET-4 Test Syllabus and Sample Test Paper (Revised Version)*. Shanghai: Shanghai Foreign Language Education Press.
- Syllabus for College English Test. (2006). National College English Testing Committee, Shanghai, China: Shanghai Language Education Press.
- Yang, H. (2000). Validation study of the College English Test. Paper presented at the International Association of Applied Linguistics (AILA). Waseda, Japan.
- Yang, H. & Jin, Y. (2001). Score interpretation of CET. *Foreign Language World* 81, 62–68.
- Yang, H. & Weir, C.J. (1998). *The CET validation study*. Shanghai: Shanghai Foreign Language Education Press.
- Zhang, X. (2003). On the negative impact of College English Test Band 4. *Educational Science* 19, 35–39.