

# Chapter 4

## Washback of the Reformed College English Test Band 4 (CET-4) in English Learning and Teaching in China, and Possible Solutions



Feifei Han

**Abstract** This chapter reviews the key research on the washback of the College English Test Band 4 Test (CET-4), a compulsory high-stakes, large-scale, and nationwide English proficiency test to measure English proficiency of non-English major university students in China. The review results show a mixture of positive and negative effects of washback on both English learning and teaching. To reduce the negative effects of the washback, the following solutions are proposed: (1) The quality assurance body of College English teaching should use a wider range of indicators to gauge the quality of English teaching and learning rather than solely relying on the outcomes of the CET-4. (2) The Spoken English Test should be designed as a compulsory subset so that the English-speaking skill would no longer be devalued and ignored in College English teaching and learning. (3) The proportion of the Chinese to English translation section should be decreased to discourage using rote memorization of bilingual vocabulary lists as a main test preparation strategy. (4) The CET-4 should use the integrated format to replace the separate testing of listening, reading, and writing so that communicative English competence can be effectively assessed.

### 4.1 Introduction: Purpose and the Testing Context of the CET-4

China has the largest number of English language learners in the world. It is estimated that more than 400 million Chinese learn English as a foreign language across the nation (Wei and Su 2012). Among them, around one-tenth of the learners are university students, and College English is a compulsory subject for all freshmen and sophomores in China (Cheng and Curtis 2010). Chinese university students take a variety of high-stakes international and national English tests to satisfy their different purposes and needs. In order to study overseas in an English-speaking country, they

---

F. Han (✉)  
Griffith University, Brisbane, Australia  
e-mail: [feifei.han@griffith.edu.au](mailto:feifei.han@griffith.edu.au)

take TOEFL, IELTS, and the Pearson Test of English. In order to be competitive in job markets, they take TOEIC and Cambridge English Qualifications: Business. Apart from these international English tests, the most important large-scale standardized English test for Chinese university students at the national level is the College English Test (CET). The CET has two levels: level 4 (known as CET-4) and level 6 (known as CET-6), both of which are administered by the Ministry of Education in China (CMoE). While both tests examine the English proficiency of Chinese non-English major university students, CET-4 is compulsory, whereas CET-6 is not (Yang and Weir 1998). The compulsory nature of the CET-4 renders it to have high-stakes status, with around 10 million test takers annually (Zheng and Cheng 2008). CET-4 is developed by the National College English Testing Committee (NCETC) and is held twice a year in June and December. Currently, the test is open only to the currently enrolled undergraduates, and the registration of the test must be completed through the test takers' universities.

First launched in 1987 following the implementation of the first College English Curriculum (CMoE 1985, 1986), the initial development of the CET-4 aimed to achieve two goals. The first goal was to provide an objective assessment and evaluation of a university student's overall English proficiency. The second goal was to direct and to unify the College English teaching nationwide (Ma 2014).

Throughout its 31 years of administration and development, the CET-4 has undergone two waves of major reforms in terms of content, format, and scoring system (Jin 2017). The first wave of reform was implemented throughout the 1990s, and this wave involved three noticeable changes. The initial test items in the CET-4 were predominantly multiple-choice questions, which accounted for as much as 85% of the test items, and only 15% were essay writing for testing English writing skills. Although having a high percentage of multiple-choice questions has potential advantages in terms of objectivity in testing and efficiency for scoring for a large-scale test, and are appropriate for testing receptive skills, such as listening and reading skills, and the receptive knowledge of vocabulary and grammar, this format has apparent drawbacks for evaluating productive skills (Jin and Wu 2017). To solve this issue, the first change of the CET-4 was the inclusion of a variety of formats, including dictation of short English phrases and sentences, short answer questions, and translation from English to Chinese. The second major change concerned the scoring, with the revised scoring system emphasizing writing. A threshold was set for the writing subset, and not achieving the threshold in writing resulted in a penalty of the total score of the CET-4. If a student scored zero for writing, he/she had to retake the whole test, no matter what level of performance he/she had achieved for the remainder of the test. The third major change was the introduction of a spoken test in 1999, known as CET Spoken English Test (CET-SET) to assess students' English-speaking skills (Zhang 2005). However, the CET-SET was not open to all candidates, as the test takers had to achieve an overall 80 out of 100 points in the test to be eligible for the CET-SET. The CET-SET had three-level grades: A, B, and C. Grades lower than C did not produce a report, as this indicated not having a sufficient level of English-speaking skills.

The second wave of reform of the CET-4 was launched in 2006 to meet the needs of the transformation of the 1999 version to the 2007 version of the National College

English Teaching Curriculum and Syllabus and Requirements (CMoE 1999, 2007) as part of the large-scale project—Higher Education Undergraduate Level Teaching Quality and Teaching Reform by the CMoE (2006). The wave of the reform had two major purposes: (1) to meet “the pressing social need for college and university graduates with a stronger communicative competence in English” (Jin and Yang 2006, p. 21), and (2) to “maximize its positive backwash effect on teaching and beneficial impact on society” (p. 34). The reform was implemented by the NCETC, which constructed new tasks to replace some old ones in order to test contextualized English use, rather than the context-free of English language knowledge (NCETC 2006).

The reform included four major changes: (1) increasing the proportion of testing listening comprehension from 20 to 35%; (2) replacing the multiple-choice style of assessing English vocabulary and structure in single sentences with a contextualized cloze test; (3) adding fast reading to assess learners’ skimming and scanning abilities in reading longer English texts; and (4) replacing translation from English to Chinese with translation from Chinese to English. The structure of the reformed CET-4, including subsets/skills, contents, formats, proportions, and time distribution, is displayed in Table 4.1.

Apart from the revision of the test format, the scoring system has also been dramatically changed. In the previous scoring system, the test takers only received a certificate indicating if they achieved a pass (60) or distinction (85) out of 100 points. The new scoring system is norm-referenced; hence the scores show how an individual test taker has performed relative to the whole group. The maximum achievable score is 710 points, and the test takers receive a report of the total score and the scores of the subsets of listening (maximum achievable score is 248.5), reading (maximum achievable score is 248.5), writing (maximum achievable score is 106.5),

**Table 4.1** The structure of the reformed CET-4

Subsets	Contents	Formats	Number of questions	Proportions (%)	Time distribution
Listening	3 short news	multiple-choice	7	7	25 minutes
	2 long conversations	multiple-choice	8	8	
	3 short texts	multiple-choice	10	20	
Reading	vocabulary	cloze test	10	5	40 minutes
	fast reading	information matching	10	10	
	in-depth reading	multiple-choice	10	20	
Translation	Chinese to English	paragraph translation	1	15	30 minutes
Writing	composition	short essay	1	15	30 minutes
Total	—	—	57	100	125 minutes

and translation (maximum achievable score is 106.5). The new scoring system has also removed “pass” to prevent universities from using the CET-4 pass as a compulsory requirement for students to graduate and to prevent employers from using it as a mandatory selection criterion for job interviews.

These strategies aimed at reducing the pressure on the test takers and English teachers in the hope that the test can induce positive washback effects in College English learning and teaching (Jin 2010). Therefore, since the introduction of the second reformed CET-4, a few studies have been conducted to examine if the test has achieved its goal of stimulating positive effects. This review only covers the key studies on the washback of the reformed CET-4 (after 2006). The findings of the key studies are summarized, and the problems in terms of negative washback are outlined in the next section.

## 4.2 Testing Problems Encountered

In the language testing literature, washback is defined as “the effect of testing on teaching and learning” (Hughes 2003, p. 1). A test is able to produce both positive and negative washback on teaching and learning (Bachman and Palmer 2010). A search of the literature identified nine key studies, whose results showed a mixture of positive and negative washback of the CET-4 on Chinese students’ English learning and the instruction of College English by teachers. This review excluded the studies which only focused on investigating the washback of one of the subsets tested in the CET-4 (e.g., the listening subset: Hou and Wang [2008], Shi [2010], Wang [2010]; the speaking subset: Zhuo [2017]). The results of the nine key studies can be broadly categorized as the washback of the CET-4 on learning and on teaching. For each category, both positive and negative effects were observed and will be discussed in turn. A detailed summary of the nine key studies, including types of research, participants and data collection methods, and key findings, is presented in Appendix 4.1.

Concerning the washback effect on English learning, the reformed CET-4 produced three major positive effects. First, the CET-4 test not only enabled students to put much time and effort in preparing for it, but also motivated students to make extra efforts in learning English (i.e., Li et al. 2012; Shao 2006; Sun 2016). The survey in these three studies out of the nine key studies reported that the majority of students felt the positive side of having to sit for the CET-4 was that they were motivated to learn English. Second, apart from putting much effort into English learning, some students also reported that they felt CET-4 made them more aware of the goals of English learning and kept them focused on the goals (i.e., Li et al. 2012). Third, CET-4 encouraged students to use cognitive strategies and test management strategies in the test (i.e., Xiao 2014).

There were also three prominent negative effects of the CET-4 on English learning. To start with, in five out of the nine key studies, the results showed that students adopted English learning strategies oriented toward passing the test rather than

strategies enhancing language use competence and developing English communicative abilities, in the processes of test preparation (i.e., Ren 2011; Sun 2016; Xiao 2014; Xie and Andrews 2012; Zhan and Andrews 2014). These test preparation strategies encouraged rote memorization of linguistic forms, such as memorizing English–Chinese bilingual vocabulary lists; using grammar-translation methods to learn English and to prepare for the test; and developing test-wise skills through practicing the mock or past CET-4 papers, by means such as analyzing test papers, rehearsing test-taking strategies, practicing sample test papers intensely, and memorizing the model English essays. Students also used past tests and sample CET-4 papers as their main sources to learn English. Furthermore, students put much effort into and gave much weight to practicing reading and listening, which were two major skills tested in CET-4. On the other hand, they neglected writing and speaking, because writing only accounted for a small proportion and speaking was not compulsory (i.e., Li et al. 2012). Last, apart from the negative effects on cognition, CET-4 also produced psychological effects detrimental to students, as the test aroused pressure and anxiety in their English learning (i.e., Li et al. 2012).

In terms of the washback on English teaching and teachers, three major positive and negative effects were identified. In terms of positive washback on teaching, firstly, a greater importance was attached to English teaching due to the compulsory nature of CET-4 (i.e., Gu 2007). Secondly, CET-4 promoted the implementation of the new version of the National College English Teaching Curriculum and Syllabus and Requirements in the College English instruction (i.e., Gu 2007). Thirdly, English teaching shifted from grammar drilling to developing communicative competence in English, and teachers tended to integrate the four skills in their instruction (i.e., Chen 2007).

The first major negative impact of CET-4 on English teaching was its narrowing of the content of instruction (i.e., Gu 2007; Ren 2011; Shao 2006). The teaching strategies, teaching materials (e.g., the sample test papers are exclusively used in teaching), and teaching activities (e.g., teachers only practice the skills tested in CET-4 but ignore the skills not tested in CET-4, such as spoken English) tended to be exclusively focused on when CET-4 was approaching. As a result, teachers were unwilling to develop students' communicative competence, and the teaching lacked creativity. In the time period before CET-4, the normal schedule of English teaching was always disrupted, and teachers found it difficult to follow the syllabus and to complete the contents in the English textbooks.

Furthermore, CET-4 also negatively affected the content of classroom assessment (i.e., Ren 2011). The assessment tasks were predominantly designed to resemble those in CET-4, resulting in a lack of formative assessment of English learning. Last but not least, CET-4 cast a negative influence on teachers and teaching (Chen 2007; Gu 2007; Ren 2011). The results of students' performance on CET-4 were used heavily to benchmark the quality of English teaching and were given much weight for English teachers' promotion, which created considerable pressure for teachers.

### 4.3 Solution/Resolution of the Problems

Some possible solutions are proposed in order to mitigate the negative washback produced by the CET-4. (1) The quality assurance body of the universities should use a wider range of indicators to gauge the quality of College English teaching and learning rather than solely relying on the outcomes of the CET-4. (2) The CET-SET should be integrated as a compulsory subset for all the test takers so that English-speaking skills are no longer devalued in College English teaching and learning. (3) The weights of some subsets of the tests should be modified. In particular, the writing subset should be given more weight, and the proportion of the Chinese to English translation section should be reduced to discourage using rote memorization of vocabulary lists. (4) The NCETC should consider developing a test format which integrates the assessment of the four skills, so that communicative competence can be emphasized in English learning and teaching.

### 4.4 Insights Gained

To mitigate the negative effects generated by CET-4, the National College English Testing Committee (NCETC) may need to consider the following aspects. Most importantly, due to the importance of communicative competence in English learning, testing the speaking and listening proficiency of learners cannot simply be ignored. While the current CET-SET is optional, how to properly integrate it into the CET-4 and make it compulsory remains a critical issue for the NCETC to solve in order for the speaking and listening skills to be appropriately assessed. Accordingly, such implementation may direct the curriculum of college English teaching to place much more emphasis on training students' communicative competence rather than predominantly focusing on reading and writing per se. Secondly, since the current use of the CET-4 results are closely linked with benchmarking teaching quality, the contents and formats of teaching tend to be greatly impacted by the contents and format of CET-4. Hence, the National Educational Examinations Authority should consider making a policy to regulate how the test results should be used in the universities to minimize the negative influence of CET-4 on normal English teaching.

### 4.5 Conclusion: Implications for Test Users

This review indicates that there is a lack of nationwide research on the washback of the reformed CET-4 since Gu (2007). Considering the high-stakes status of the test, it is suggested that the NCETC should conduct large-scale studies on the washback of CET-4 so that the intended use, formats and contents of the test may be modified to address negative effects. The NCETC should also make the interpretation of the test

scores transparent by possibly including qualitative and descriptive references as to what a test taker “can do” in English and/or the subskills of English for a particular range of scores and/or sub-scores by a test taker (Jin 2011).

This review also has important implications for stakeholders of the CET-4, including government, employers, and universities. It is risky to rely solely on the scores of the CET-4 as a means to gauge English learning and teaching. Students and teachers should develop strategies to develop English competence in the long run, which may in turn facilitate achieving a desirable level of performance in CET-4.

### Appendix 4.1: Summary of the Key Studies on the Washback of the Reformed CET-4

Studies	Research method	Participants and data collection methods	Positive effects	Negative effects
Shao (2006)	mixed-methods	45 teachers from three teachers' colleges were surveyed 356 students from the three teachers' colleges were surveyed Of 45 teachers, four were interviewed Each interviewed teacher was observed for four hours of English teaching	70% of the students reported that they were motivated to learn English due to the CET-4	The CET-4 did not exert much influence on freshmen However, the CET-4 produced negative effect for sophomores, who would sit the CET-4 soon The teaching strategies, teaching materials (i.e., The sample papers were exclusively used in teaching), and teaching activities (e.g., Only practicing the skills tested in the CET-4 but ignoring the skills which were not tested in the CET-4, such as the spoken English) were exclusively focused on the CET-4, as the test was approaching

(continued)

(continued)

Studies	Research method	Participants and data collection methods	Positive effects	Negative effects
Chen (2007)	mixed-methods	154 teachers were surveyed Of the surveyed teachers, 112 were interviewed	Teachers felt that the revised CET-4 had positive effects on their curricular planning and instruction: (1) 100% of teachers reported that they were motivated to integrate listening and speaking skills in the teaching activities rather than only targeted reading and writing as in their previous teaching (2) Over 87% teachers reported that the teaching focus was shifted from grammar drilling to developing communicative competence	A positive correlation was found between teachers' perceptions of the importance of the CET-4 and the perceptions of the pressure in their teaching
Gu (2007)	mixed-methods	2609 students nationwide were surveyed 1220 teachers nationwide were surveyed English teaching of 38 teachers at three universities was observed Focus-group interviews were conducted with teachers, students, and administrators	(1) The majority of the teachers, students, and administrators felt positively about the CET-4 because the university attached greater importance to English teaching and learning due to the compulsory nature of the test (2) The CET-4 also promoted the implementation of the new version of the <i>National College English Teaching Curriculum and Syllabus and Requirements</i>	(1) When the CET-4 approached, teachers used the CET-4 preparation materials predominantly in teaching. As a result, they could not cover all the contents in the English textbooks, and the pace of English teaching was also accelerated (2) The administrators used the CET-4 results as the sole indicator to evaluate the quality of English teaching quality and students' English proficiency

(continued)



(continued)

Studies	Research method	Participants and data collection methods	Positive effects	Negative effects
Ren (2011)	mixed-methods	35 teachers were surveyed 210 students were surveyed Among 35 teachers, 22 were interviewed Among 210 students, 30 were interviewed		(1) The CET-4 encouraged rote memorization of linguistic forms (e.g., memorizing vocabulary lists) and practicing past CET-4 papers to prepare for the test. As a result, students were unable to use English in authentic situations (2) Classroom assessments highly resembled the formats and the contents of the CET-4 (3) Teachers were unwilling to train students' communicative competence in English (4) Teachers' promotion was linked to the CET-4 success rate achieved by their students (5) It suppressed teachers' creativity in teaching
Li et al. (2012)	quantitative	150 students at a university were surveyed	Most of the students surveyed felt the CET motivated them to invest greater effort to learn English (2) The CET-4 enabled students to set a clearer goal for English learning	(1) The students felt pressure and anxiety in English learning (2) The students put much effort in practicing reading and listening skills, which were given more weight in the CET-4, and they tended to neglect the writing and speaking skills, as the two skills were either given little weight or not tested

(continued)

(continued)

Studies	Research method	Participants and data collection methods	Positive effects	Negative effects
Xiao (2014)	quantitative	284 students were surveyed	The CET-4 moderately promoted cognitive strategy use and weakly promoted test management strategy use	The students' English learning strategies were test-oriented (e.g., using test-wise strategies) rather than developing competence of language use
Xie and Andrews (2012)	quantitative	870 students were surveyed		The students' perceptions of test design and test use affected their test preparation strategies, which were dominated by analyzing past test papers, rehearsing test-taking skills, practicing sample test papers intensively, and memorizing model writing essays in a rote manner
Zhan and Andrews (2014)	qualitative	24 students were required to keep diaries and were also interviewed following the diary entries		<p>(1) The students favored learning strategies of rote memorization over developing their English communicative competence</p> <p>(2) The students used the past CET-4 papers as the only learning materials</p>
Sun (2016)	mixed-methods	A CET-4 test developer from the development committee was interviewed Eight CET-4 test users from educational and social contexts (two deans of the two universities, two directors of the government employment offices; and four human resources managers of four companies) were interviewed 416 students from two universities were surveyed and they also took a retired CET-4	The students' perceptions of the high demands of the CET-4 led them to spend more effort in preparing for it	The student had high perception of instrumental uses of the CET-4. Such perception had stronger association with use of rehearsing and cramming strategy (e.g., memorizing model essays to prepare for the CET-4 writing) than use of long-term skill development strategies

## References

- Bachman, L., & Palmer, A. (2010). *Language assessment in practice: Developing language assessments and justifying their use in the real world*. Oxford: Oxford University Press.
- Chen, F. (2007). The washback effect of College English Test Band 4 on curricular planning and instruction. *CELEA Journal*, 30(1), 19–29.
- Cheng, L., & Curtis, A. (2010). The realities of English language assessment and the Chinese learner in China and beyond: English language assessment and the Chinese learner. In L. Cheng & A. Curtis (Eds.), *English language assessment and the Chinese learner* (pp. 3–12). New York: Routledge.
- CMoE. (1985). *National College English curriculum syllabus for science and technology students*. Shanghai: Foreign Language Education Press.
- CMoE. (1986). *National College English curriculum syllabus for liberal arts students*. Shanghai: Foreign Language Education Press.
- CMoE. (1999). *College English curriculum syllabus* (revised version). Shanghai: Shanghai Foreign Language Education Press.
- CMoE. (2006). *Quality and reform of college English education*. [http://www.moe.gov.cn/publicfiles/business/htmlfiles/moe/s3857/201011/xxgk\\_110823.html](http://www.moe.gov.cn/publicfiles/business/htmlfiles/moe/s3857/201011/xxgk_110823.html). Accessed 13 Feb 2019.
- CMoE. (2007). *College English curriculum requirements*. [http://old.moe.gov.cn/publicfiles/business/htmlfiles/moe/moe\\_1846/200711/28924.html](http://old.moe.gov.cn/publicfiles/business/htmlfiles/moe/moe_1846/200711/28924.html). Accessed 13 Feb 2019.
- Gu, X. (2007). *Positive or negative: An empirical study of CET washback*. Chongqing: Chongqing University Press.
- Hou, X., & Wang, W. (2008). A study of the college students' attitudes towards the new CET-4 listening subset: Washback effect of the new CET-4 listening subtest on students. *Journal of Xi'an International Studies University*, 16(3), 91–94.
- Hughes, A. (2003). *Testing for language teachers* (2nd ed.). Cambridge: Cambridge University Press.
- Jin, Y. (2010). The National College English Testing Committee. In L. Cheng & A. Curtis (Eds.), *English language assessment and the Chinese learner* (pp. 44–59). New York: Routledge.
- Jin, Y. (2011). Fundamental concerns in high-stakes language testing: The case of the College English Test. *Journal of Pan-Pacific Association of Applied Linguistics*, 15(2), 71–83.
- Jin, Y. (2017). Construct and content in context: Implications for language learning, teaching and assessment in China. *Language Testing in Asia*, 7(1). <https://doi.org/10.1186/s40468-017-0044-1>.
- Jin, Y., & Wu, E. (2017). An argument-based approach to test fairness: The case of multiple-form equating in the College English Test. *International Journal of Computer-Assisted Language Learning and Teaching*, 7(3), 58–72.
- Jin, Y., & Yang, H. (2006). The English proficiency of college and university students in China: As reflected in the CET. *Language, Culture and Curriculum*, 19(1), 21–36.
- Li, H., Zhong, Q., & Suen, H. (2012). Students' perceptions of the impact of the College English Test. *Language Testing in Asia*, 2(3), 77–94.
- Ma, F. (2014). College English Test: To be abolished or to be polished. *Journal of Language Teaching and Research*, 5(5), 1176–1184.
- NCETC. (2006). *CET-4 Test syllabus and sample test paper* (2006 revised version). Shanghai: Shanghai Foreign Language Education Press.
- Ren, Y. (2011). A study of washback effects of College English Test (Band 4) on teaching and learning English at tertiary level in China. *International Journal of Pedagogies and Learning*, 6(3), 243–259.
- Shao, H. (2006). An empirical study of washback from CET-4 on college English teaching and learning. *CELEA Journal*, 29(1), 54–59.
- Shi, X. (2010). A longitudinal study on the washback effect of the new CET-4 and CET-6 listening subset. *Foreign Language World*, 177(3), 80–86.
- Sun, Y. (2016). *Context, construct, and consequences: Washback of the College English Test in China* (Unpublished doctoral dissertation). Queens' University, Kingston, Canada.

- Wang, W. (2010). Investigating the washback effect of the new CET-4 listening comprehension subtest on language learners. *Chinese Journal of Applied Linguistics*, 33, 28–39.
- Wei, R., & Su, J. (2012). The statistics of English in China. *English Today*, 28(3), 10–14.
- Xiao, W. (2014). The intensity and direction of CET washback on Chinese college students' test-taking strategy use. *Theory and Practice in Language Studies*, 4(6), 1171–1177.
- Xie, Q., & Andrews, S. (2012). Do test design and uses influence test preparation: Testing a model of washback with Structural Equation Modeling. *Language Testing*, 30(1), 49–70.
- Yang, H., & Weir, C. J. (1998). *The CET validation study*. Shanghai: Shanghai Foreign Language Education Press.
- Zhan, Y., & Andrews, S. (2014). Washback effects from a high-stakes examination on out-of-class English learning: Insights from possible self-theories. *Assessment in Education: Principles, Policy & Practice*, 21(1), 71–89.
- Zhang, Z. (2005). A historical review of CET 4/6. *Modern Language Journal*, 18(2), 100–103.
- Zheng, Y., & Cheng, L. (2008). College English Test (CET) in China. *Language Testing*, 25(3), 408–417.
- Zhuo, W. (2017). The Washback effect of CET Spoken English Test upon college English teaching. *Canada Social Science*, 13(1), 62–68.