

## The CET-4: Encouragement or Impediment of Undergraduate English Teaching and Learning

**Xi Chen**

Harbin Institute of Technology  
Shenzhen, China

**Joyce Gail Webb**

Foreign Expert in Communications  
Harbin Institute of Technology  
Shenzhen, China

### Abstract

---

*This paper constitutes a review of the literature pertaining to CET-4, the College English Test Band 4, a national English test designed to examine the English proficiency for non-English majors in colleges and universities across China. Through literature review, the focus of the paper is to investigate whether the test is viewed as an encouragement or impediment to undergraduate English teaching and learning. The backgrounds of positive and negative washback associated with CET-4 are discussed. Scopes, theories and models are reviewed. Particularly, a local situation in Harbin Institute of Technology (Shenzhen) is included to see the influence of CET-4 from teachers and students. Though causing some problems with negative washback, CET-4, as a test to evaluate students' English proficiency, has proved to be valuable for both teaching and learning.*

---

**Keywords:** Washback, CET-4, China, College and University, English Teaching and Learning.

### Introduction

The CET-4, College English Test Band 4, a national English test for non-English majors, is designed to examine the English proficiency of undergraduate students to check whether they have met the requirements of the National College English Teaching Syllabi (NCETS). Aiming to promote the enforcement of the NCETS, the test is expected to foster English teaching and learning. Considered as the most crucial English test of undergraduate students, this test certificate was somehow linked to the successful attainment of a degree and later the guarantee of a well-paid job. In some universities, the results of students' CET-4 are closely related to the English course curriculum and the assessment of English teachers. Therefore, some teachers would adopt test-oriented teaching to ensure the scores of their students. In doing so, teachers could be torn between focusing on teaching examination skills or teaching valuable knowledge and skills including: western culture, problem-solving skills or critical thinking. Problems also occur with students in the classrooms. Often, some of the keen test-takers are not willing to listen to what the teachers say in class. Instead, they will take either a CET-4 vocabulary book to memorize words by themselves or bury themselves in a pile of CET-4 model tests to finish them in class. Considering the problems that CET-4 brings in college English teaching and learning, it would be prudent to investigate, further, whether the test is an encouragement or impediment for teachers and students. The analysis will include: Background, Literature Review, Discussion, Limitation, Conclusion and Implication.

## **Background of CET-4**

The CET-4 was first introduced in 1987 by the Department of Higher Education, Ministry of Education (MOE) in China. It has passed its third decade of prevalence. (Yang, 2003; Jin, 2004) From its first test until now, it has dominated higher education institutes in China, which means almost all the Chinese college students have the opportunity to take the exams. Designated by the MOE, the National College English Testing Committee is authorized to hold the CETs twice a year to give an objective, impartial, and accurate assessment of the English proficiency for college students in comparison with the desideratum in NCETS. To compensate for the lack of an oral test, the CET-SET (College English Test-Spoken English Test) was carried out in 1999 to improve the assessment system of students' English proficiency. (Yang, 2003; Jin, 2004). As China moves from Reform and Opening-up to later expansion of the recruitment in higher education as well as many other international participations, like entering the WTO and holding the 2008 Olympic Games, the CET test-takers are experiencing an exponential increase. According to the data provided by Professor Yan Huizhong, the number of CET-4 test takers was about 100 thousands in 1987. By the year 2002, it has increased to more than 4 million students. (2003: 21). Consequently, there is the nationwide acceptance and recognition of CETs not only in higher education but also in the workplace. (Yang, 2003:23).

## **Contribution of CETs**

Advocates for the test often claim that the CETs are established to be a support of college English teaching and learning. With this test, colleges and universities place considerable emphasis on the English classroom. Both English teachers and students are highly motivated to take these tests to meet the standards of the NCETS. The test results would provide beneficial feedback to teachers and students to improve English teaching and learning. As for society, the tests have also served as a standardized qualification for employers to evaluate their candidates' English proficiency. (Yang, 2003)

## **Negative Impacts of CETs**

With its unquestionable authority and widespread acknowledgement, CETs, as high-stake tests are becoming utilitarian in contemporary society. To compete for the ranking of CETs, some colleges and universities have set up the "linked policy" to connect the CET certificate and the degree diploma together. (Yu, 1999:33; Yang, 2003:25) In these circumstances, both teachers and students are highly test-concentrated and under great pressure to do well in the CETs. Accordingly, English teachers would usually abandon teaching for improving English proficiency. Instead, they concentrate on examination skills in class. (Yu, 1999; Yang, 2003) In some schools, problems labelled as the "sea tactics", which is doing piles of CET-4 coached materials and using the "spoon-feeding method" of teachers forcing students to learn, often occur. (Yu, 1999: 33; Yang, 2003:25) As a result, listening, reading and writing are heavily stressed, while speaking is not the focus because the CET-SET is not a compulsory test for all the students. With the purpose of getting their degrees and later gaining access to a proper job, the students are obsessed with the accomplishment in CETs. (Yang, 2003) However, it also gives rise to a series of problems, such as producing those with high scores and low abilities, and students who can merely speak dumb English. (Yu, 1999)

## **Literature Review**

### **Washback: Definitions**

In language testing, the term "washback" is referred to as the influence or impact a test can have not only on teachers but also on students, either positively or negatively, in language teaching and learning. (Liu and Gu, 2013)

In the academic field, scholars and researchers have summed up their definitions and expressions concerning "washback".

In a paper regarding testing listening comprehension in Japanese university entrance exams, Gary Buck (1988:17) has elaborated that testing has significant effects on classroom teaching. With the demand for the test, it is natural that both teachers and students tailor the classroom to meet the requirements. This is the case even "when the test is very important to the future of the students, and pass rates are used as a measure of teacher success". Washback, the influence of the test on the classroom can be "either beneficial or harmful". (1988:17)

Shohamy (1992:513) develops the term “washback effect” or “measurement-driven instruction” in a radical way. She explains that “language tests that have been traditionally used to make decisions and judgments about the future of individuals are now being used for influencing the educational context of learning.” Consequently, the authority of these tests has a vital impact on the life of the test takers. The tests have been so powerful that they can “change and prescribe the behaviour of those administrators, teachers, and students who are affected by their results”. Endowed with such power, “central agencies and decision makers have often used (or abused) external tests to impose new curricula, new text books, and new teaching methods. Thus, external tests are currently used to force students to study, teachers to teach, and principals to modify the curriculum”.

Alderson and Wall (1993:117) indicate that with the impact of the washback, “teachers and learners do things they would not necessarily do because of the test”. This viewpoint reiterates the idea of Shohamy that tests could drive people-administrators, teachers, and students to do things they don't need to do.

Extending Alderson and Wall's notion, Messick (1996:241) specifies that washback is “the introduction and use of a test which influences language teachers and learners to do things they would not otherwise do that promote or inhibit language learning”. This means the washback effect cuts both ways.

According to Bachman and Palmer (1996:30), washback is viewed as the “direct impact of testing on individuals and it is widely assumed to exist”. Concentrating on the learning and instruction process which takes place and is in practice, tests have “effects on the individuals, educational systems and society”. Washback should be best taken into consideration “within the scope of impact”.

To conclude, the definitions provided by the above scholars and researchers indicate that washback means a test can have either beneficial or harmful impact on individuals, educational systems and society. When the test has high stakes, it could trigger harmful impacts to force people to do things that could thwart the original purpose of testing.

## **Classification of Washback**

### **Positive and Negative Washback**

One of the classifications of washback is that it is positive and negative, as it is presented in the definition.

To distinguish positive and negative washback, we should judge the effects of examinations on what we do in the classroom. (Hughes 1989, Heaton and Han 2012) Simply put, if a test can encourage teaching and learning, it will produce positive washback. On the contrary, when teaching and learning are impeded, it will exert negative washback.

To shed light on the positive washback, Fredericken and Collins (1989) form a notion called “systemic validity” of a test, which considers the effects of instructional changes that a test could bring into the education system. The positive connotation embodies the idea that a test which prompts the changes in curricular and educational systems stimulates the “development of the cognitive skills that the test is designed to measure” (27).

In Hughes's book *Testing for Language Teachers* (1989), one chapter has been devoted to analyzing the ways of achieving beneficial washback. Among the seven ways of developing positive washback are:

- 1) Test the abilities whose development you want to encourage.
- 2) Sample widely and unpredictably.
- 3) Use direct testing.
- 4) Make testing criterion-referenced.
- 5) Base achievement tests on objectives.
- 6) Ensure [that the] test is known and understood by students and teachers.
- 7) Where necessary provide assistance to teachers. (44-47)

Based on Hughes, Bailey (1996: 268-272) four ways of promoting beneficial washback are summarized.

- 1) language learning goals;
- 2) authenticity;
- 3) learner autonomy and self-assessment;
- 4) detailed score reporting

As for negative washback, Vernon (1956: 166) has asserted that examinations “distort curriculum” and teachers tend to neglect subjects and activities that are not helpful to pass the exams.

The negative connotations of washback also refer to “measurement-driven instruction” and “test-curriculum alignment”. “Measurement-driven instruction” means that instruction is driven by the tests. And “test-curriculum alignment” reveals the relationship between test contents and curriculum, showing that the curriculum could be narrowed by the teaching of the test. (Chen, 2007)

Alderson and Hamp-Lyons present the four major factors that could induce negative washback.

- 1) Narrowing of the curriculum
- 2) Lost instructional time
- 3) Reduced emphasis on skills that require complex thinking or problem-solving
- 4) Test score ‘pollution’, or increases in test scores without an accompanying rise in ability, in the construct being tested (1996:281).

### Overt and Covert Washback

Another categorization is overt and covert washback. Raised by Prodromou (1995:14-15), overt washback is usually negative and explicitly identifiable. It includes “doing a lot of past papers in class as preparation for an examination; replicating from past papers or the textbook, the exercise types favoured in the particular examination; students taking multiple-choice, transformation, or gap-filling”. In this case, teachers and students will gradually become exam-oriented and test-driven. As for the teaching and learning contents, they will be exclusively focused on what is related to the test, ignoring other knowledge and skills.

The counterpart of overt washback is covert washback, being known as the implicit, deep-seated and often unconscious process. In accordance with Prodromou (1995), the covert testing is more elusive and disturbing because teachers regard teaching a textbook as teaching a series of tests. Unaware of this, many teachers conduct the class like this: question a student, obtain an accurate answer from the student, get another question for another student, not engaging the rest of the class. It, therefore, “makes the activity more of an informal assessment than a teaching procedure” (15).

### Theoretical Background of Washback

The past two decades have witnessed the enrichment of discussion on the scope, pattern and theoretical frameworks of washback, from general to specific. Among them, we will demonstrate the following four frameworks:

#### Bachman and Palmer’s Pattern

Bachman and Palmer (1996: 29-30) develop washback into two levels. They conceptualize the impact of a test on a micro level-the influence of a test could have on individuals and a macro level-the impact of the test on educational systems and the society.

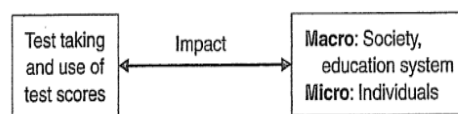


Figure 1: Impact of a test

They consider the impact of the test comes from the test itself and the use of test results. However, it is deeply rooted in the social and educational system, because it has empowered the test with specific values and goals. Therefore, the test, in turn, will extend influence on the society, the educational system (macro level) and individuals (micro level). This pattern not only illustrates the scope and level of test impact but also reveals the dynamic and complicated relationships between the test and society, educational systems and individuals.

### Alderson and Wall's Washback Hypothesis

By proposing 15 questions in the paper "Does washback exist?", Alderson and Wall (1993: 120-121) explore some possible hypotheses.

- 1) A test will influence teaching.
- 2) A test will influence learning.
- 3) A test will influence what teachers teach.
- 4) A test will influence how teachers teach.
- 5) A test will influence what learners learn.
- 6) A test will influence how learners learn.
- 7) A test will influence the rate and sequence of teaching.
- 8) A test will influence the rate and sequence of learning.
- 9) A test will influence the degree and depth of teaching.
- 10) A test will influence the degree and depth of learning.
- 11) A test will influence attitudes to the content, method, etc., of teaching and learning.
- 12) Tests that have important consequences will have washback.
- 13) Conversely, tests that do not have important consequences will have no washback.
- 14) Tests will have washback on all learners and teachers.
- 15) Tests will have washback effects for some learners and some teachers, but not for others.

To develop the washback hypotheses further, Alderson and Hamp-Lyons (1996: 296) extend one more of the above.

- 16) Tests will have different amounts and types of washback on some teachers and learners than on other teachers and learners.

Qi Luxia (2004) classifies the fifteen hypotheses into three groups. Group one (1, 3, 4, 7, 9, and 11) is washback on teaching, which concentrates on the teaching aspects of contents, methods and attitudes. Group two (2, 5, 6, 8, 10, 11) includes washback on learning, containing the learning facets of attitudes, content, strategies, degree and depth. Group three (12, 13, 14, and 15), strength of washback, involves varieties "from the importance of the test and the characteristics of individuals" (38).

### Hughes's PPP Model

Hughes (1993) develops the mechanisms of washback into a trichotomy idea. In his unpublished paper, he introduces the PPP model, which introduces *participants, process and products*. Participants are those whose "perceptions and attitudes towards their work may be affected by a test" (1993: 2), say, students, classroom teachers, administrators, materials developers, and publishers. Process means "any actions taken by the participants who may contribute to the process of learning" (1993: 2). Materials development, syllabus design, changes in teaching methodology, and the use of learning and/or test-taking strategies are accounted in the process. Product indicates "what is learned (facts, skills, etc.) and the quality of the learning (fluency, etc.)" (1993: 2).

Furthermore, Hughes (1993: 2) postulates:

The trichotomy into participants, process and product allows us to construct a basic model. The nature of a test may first affect the perceptions and attitudes of the participants towards their teaching and learning tasks. These perceptions and attitudes in turn may affect what the participants do in carrying out their work (process), including practicing the kind of items that are to be found in the test, which will affect the learning outcomes, the product of that work.

Based on Alderson and Wall's list of hypotheses and Hughes trichotomy idea, Bailey (1996) depicts the basic model of washback.

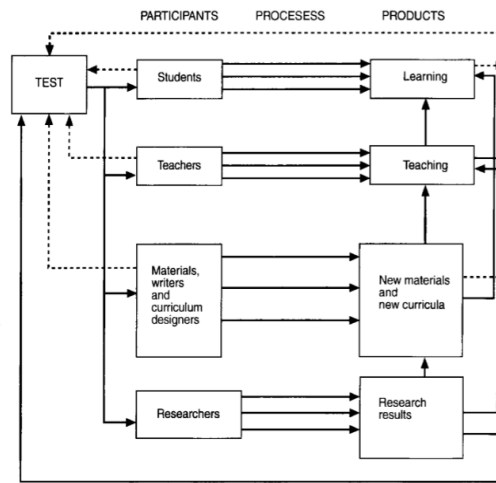


Figure 2: A basic model of washback

The solid arrow lines display the direct influence a test has on the participants, processes, and products. Possible influences from the participants on the test are manifested through the dotted lines. Bailey (1996) further elaborates the notions of “washback to the learners” and “washback to the program”, respectively “the effects of test-derived information provided to the test-takers” and “results of test-derived information provided to teachers, administrators, curriculum developers, and counsellors” (263-264).

The main contribution of Bailey is to integrate all kinds of concepts and assumptions, which reveals the complexity of washback. In her opinion, the test does not always wield direct influence on all aspects and layers of the educational system. For instance, through the test, teachers have a direct impact on teaching. Meanwhile, the test has an indirect effect on teaching through teaching materials. Therefore, the impact of the test on the various aspects of the teaching system is bilateral and even multilateral.

### Gu Xiangdong’s Model of CET Washback

To explore positive or negative CET washback, Gu Xiangdong (2005) conducts an empirical study of the CET washback on teaching and learning in her doctoral thesis. On the ground of Hughes’s PPP model, she extends it further to a basic PPPP model of CET washback.

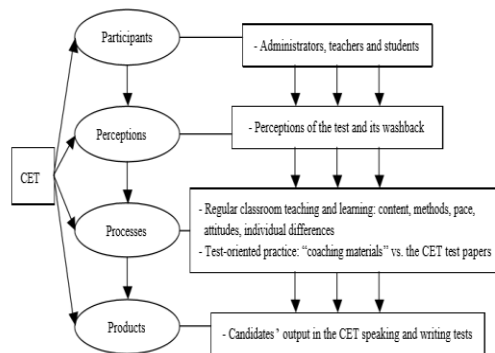


Figure 3: Basic model of CET washback (Gu Xiangdong, 2005)

As for the participants, Gu (2005) defines them as the administrators, teachers, and students who are regarded as the stakeholders of English teaching at the implementation level. Perceptions are separated as one facet to investigate participants’ views of the test and its washback on teaching and learning. Processes are put into details like test-oriented practice, regular classroom teaching and learning, including a range of elements-content, methods, attitudes and individual differences. Products are specified as the result of test-takers in the CET speaking and writing tests.

In her research, Gu (2005) concludes that evidence has been found that there are both positive and negative washback of the CET on college English (CE) teaching and learning. The positive washback includes: promoting the implementation of the *College English Teaching Syllabus* in the CE classroom; forcing administrators to attach great importance to the CE course; motivating both the teachers and students in their teaching and learning; emphasizing reading skills and reading ability. The negative washback, on the other hand, includes: “a more rapid teaching pace, use of “coaching materials” in class, and unsatisfactory completion of the textbook units in Grade Two classes, particularly in the fourth semester”. (Gu, 2005: iii)

To conclude, the above-mentioned scope, pattern, and models provide substantial contribution to the conceptualization of the washback effect. Qi (2004), summarized that Bachman and Palmer’s figure “set the scene of impact”. Hughes’s trichotomy and Bailey’s basic model of washback “elaborate the scope and pattern of impact on education systems in the scene”. Embedded in this model, Alderson and Wall’s hypotheses illustrate “the complex nature of test influence on teaching and learning”. (40-41) Entering the stage, Gu’s model of CET washback specifies, in particular, the CET washback on participants, participants’ perceptions, process, and products.

### **Empirical Studies**

As the theoretical foundation establishes, the empirical studies also flourish, both internationally and domestically. Through implementing the pattern, hypotheses and models, scholars and researchers have conducted comprehensive empirical studies, exploring different tests within various teaching and learning contexts.

### **International Empirical Studies**

In 1988, Hughes presented a report of a project conducted at Bogazici University in Istanbul, Turkey to explain the beneficial washback of a new English proficiency test on teaching and learning. Nevertheless, he fails to clarify “what washback effect the test produced, nor how it produced it”. (Alderson and Wall, 1993: 126)

In light of Chen (2007), the first empirical study published in the area of language testing comes with Alderson and Wall’s investigation on a new national examination in Sri Lanka on language teaching. The impact of the “O-level test” in a secondary school through observations and other forms of data collection, demonstrated the impact of the test on the content of teaching. However, there was no evidence of the test impact on teachers’ teaching methodology. (Wall and Alderson, 1993)

Three years later, Alderson and Hamp-Lyons (1996) undertook a study on the washback of TOEFL. By conducting interviews and observations at a specialized language institute in the USA, they compared the TOEFL preparation classes and the control groups of the non-TOEFL preparation classes. Their study showed that TOEFL had influences on the contents and methodologies of teaching. However, the impacts were variable along with teachers’ attitudes and methods of teaching. As a consequence, they supplemented Alderson and Wall’s hypothesis as mentioned above and expanded it as follows: “tests will have different amounts and types of washback on some teachers and learners than on other teachers and learners”. (296)

Focusing on the entrance examinations in Japan, Watanabe (1996, 2000, and 2001) has carried out consecutive research over years. Probing into the relationship between the university entrance examinations and the use of grammar-translation approach, he indicates that “teacher factors, such as educational background, personal beliefs and teaching experience may outweigh the possible effect of the entrance examinations”. (1996:318) In 2001, he conducted further research to identify whether the university entrance examinations stimulate the motivation of learners and to determine the effect of washback.

Shohamy and others (1996) examine the impact of two national tests in Israel: Arabic as a second language (ASL) and English as a foreign language (EFL). By means of questionnaires and interviews, the results yielded show different washback patterns for the two tests. Their research displays that “washback varies over time, owing to many factors such as the status of language and the uses of the test”. (298)

In Hong Kong, Chen (1997) presented a paper regarding the preliminary research on the washback effect of the Hong Kong Certificate of Education Examination (HKCEE), in Hong Kong secondary schools. With questionnaires, interviews and classroom observations, issues such as the nature of washback, the impact on teaching and learning, and the types of washback effect were discussed. Her study showed that the washback effect works quickly and efficiently to change the teaching contents. But, teachers are reluctant to change their teaching methods because of the constraints imposed on teachers. (38)

## Empirical Studies of CET in Mainland China

Gu Xiangdong and her research teams have conducted a series of research on the washback of CET. Gu and Peng (2010) presented a paper on probing teachers' perception of CET washback before and after CET innovation in 2005. Based on questionnaires, they drew the following conclusions: College English (CE) teachers, as a whole, are familiar with the reform. They believe that the positive impact of CET on CE teaching is far greater than the negative impact, and with the passage of time, the positive impact is on the rise. In the past, CE teaching focuses on reading and writing. So, the students' listening and speaking abilities could not be fully developed and it could lead to the problem of producing students to speak "dumb" English. The College English curriculum requirements (2007:1) pointed out that "the goal of College English teaching is to cultivate students' comprehensive ability to use English, especially listening and speaking skills." After the 2005 reform, CETs put more emphasis on the students' listening ability. According to the data provided by Gu and Peng (2010: 40), the focus of College English (CE) teaching shifts from reading to listening.

Gu, Yang and Liu (2013) investigate the longitudinal washback of the post-innovation CET on College English teaching and learning. By comparing three CE teachers' classes in 2003 with those of the same three teachers in 2009 through classroom observation, video recording, transcription, field notes, and discourse analysis, supplemented with interviews, the study indicated that the CE teaching model remains mainly unchanged before and after the CET innovation. However, clear changes and positive washback were observed in terms of rich semester teaching plans, various teaching contents, and dynamic teaching methods in 2009.

In 2014, Gu, Zhang, and Liu wrote a paper on the washback of the innovated CET since 2005 on students' extra-curricular English learning processes. This paper makes a comparative analysis on the 3 groups of students: those who prepare CET-4, those who prepare CET-6 and those who have already passed CET-6. The discussion centres on the following aspects: the types of extracurricular English learning activities, the time investment and the study purpose of the 3 groups. The results show that those students who prepare for the CET-4 and CET-6 memorize more new words and do more extracurricular English training than those who have already passed CET-6.

These results show that, after the reform, CETs have a positive role in promoting students' extracurricular English learning. But, this effect is relatively short-term. (2014:38)

Yang and Tan (2007) conducted a survey on the washback effect of CET4 writing on teaching and learning. The result of this survey showed that negative effects have outweighed positive ones. Xie (2007: IV) summarized 4 issues that raise the negative washback in CET-4: "(1) limitation of CET-4 reform; (2) rigidity and limitation of CET-4 test paper; (3) over-simplicity of CET-4's appraisal function; (4) unpopularity of CET-SET (College English Test-Spoken English Test)". Niu, Chen, and Li (2011) explore the washback of CET listening teaching based on Gu Xiangdong's 4p model. Through questionnaires and interviews with teachers, they concluded that most teachers consider the positive washback outweighs the negative one. However, the negative washback could not be neglected, which is found in the difficulty of the teaching materials, lack of the variety of accents, the exercise-stuffed method and dumb English. Sharing similar viewpoints with Xie (2007) mentioned above, Li (2011) concluded with qualitative and quantitative research that "1) the phenomena of overemphasis on the results of CET-4 and teaching and learning to test bring about a series of problems; 2) the limitation of CET-4 paper constrains CE teaching and learning and fails to really help students improve their real communicative competence". (IV) In the same year, Zhou and Liu (2011) conducted empirical research through interviews, questionnaires and observations. They proposed several issues that would raise our concerns. Some policy makers in the university decided the "linked" policy between the CET-4 certificate and the degree diploma, which put high stakes on CETs. In some policy makers' minds, the CET-4 certificate equaled the students' English proficiency. In the teachers' perspectives, if the CET-4 certificates are the decisive factor in getting a degree, the teaching contents and methods are highly restricted to the tests. In this case, the purpose of students' learning tends to be utilitarian, concentrating on getting high scores, a degree, and a well-paid job. (2011)

## Discussion

Facing the divergent perspectives, we will next discuss the CET washback from the macro aspects of society, authorities of educational systems to the micro layers of individuals like teachers and students, based on Bachman and Palmer's (1996) scope of washback impact.



### **Ministry of Education (MOE) and the 2005 Reform**

In the face of the heated debates of the positive and negative washback, an official document “The Reform Plan of CET 4 and 6” was released by Department of Higher Education, Ministry of Education in 2005. In this document, the mid-term and long-term plan of the reform was explained carefully. “After the development of the past decades, it is proved that any large-scale standardized test is a continuous improvement process. The examination committee will continue to research new test questions and study the impact of the test on teaching; at the same time, making full use of high-tech means, to achieve test online marking (CET Online Marking), to investigate Computer Based Testing (CET-CBT). The committee will focus on the study and analysis of the present language testing theory and practice, tailoring for Chinese learners of English, to more accurately test the English proficiency of college students”. (MOE, 2005)

### **Views from Authorities in the Educational System**

Professor Yang Huizhong, an expert from the National Examination Reform Supervision Committee, has acknowledged the positive impact of CETs as a larger scale, standardized test to provide objective assessment of college students. Meanwhile, he admitted that the negative impact exists in some local universities and students take the CETs as an end not a means to evaluate teaching and learning. (Yang, 2003)

Professor Jin Yang, director of the National College English Testing Committee, echoed in response to the document “The Reform Plan of CET 4 and 6” by MOE, that the CETs is accepted as an objective standardized test and will be reformed in the test questions and application of high technology in testing. (Jin, 2005)

### **Observations from HITSZ**

Next, an investigation of the CET-4 washback in the case of Harbin Institute of Technology Shenzhen (HITSZ) will be conducted to find out the washback on the micro layers of individuals like teachers and students. After comparing the above-mentioned scope, pattern, and models, Gu’s model of the CET specifically analyzes the CET washback on participants, perceptions, process, and products. Therefore, we will apply her model to our investigation.

Like the Participants, the administrators of Academic Affairs Office (AAO) in HITSZ are also under pressure to achieve good results in the CET-4. They have constantly required the English teaching team to provide help, such as lectures and workshops, for the CET-4 preparations. On one hand, it has exerted positive washback because the preparation motivates students to learn English. On the other hand, it could be harmful to the students’ English learning because all the lectures and workshops are concentrated on exam skills to get a high mark in the test. Therefore, “sea tactics” and “coaching materials” are covered. Most importantly, CET-SETs (Spoken English Test) are not compulsory. Thus, it is highly possible that students will merely focus on improving listening, reading, writing, and translation instead of practicing oral English. This could lead to the problems of “Dumb English”.

Apart from the pressure from the AAO in HITSZ, English teachers still have some freedom to design the teaching process, contents and the teaching methods. Thanks to the continual debates, reform and the regulations from MOE, HITSZ does not link the attainment of a degree with the CET-4 certificate. College English teaching is not largely restricted to the tests. For the course College Reading and Writing, the English teaching team has agreed on a unified textbook: *New Horizon College English Reading and Writing 1 (Third Edition)*. However, different teachers could develop their course according to personal interests and class variations, instead of following, strictly, the exam-oriented methods to train for CET-4. As an example in Unit 1 Section A: Towards a Bright Future for all, a more interactive method could be conducted on teaching such as holding a discussion on pre-reading, to ask students to discuss questions like “what advice their parents gave them before they came to college”. At the same time, we have time to discuss in groups some critical thinking questions like “What are you going to do at the university in order to sample widely and challenge yourself” to develop students’ logical and analytical thinking. Also, we can watch a video about a speech dedicated to college freshmen. In terms of vocabulary learning, we do not necessarily do drilling exercises but focus on the pragmatic uses of words. For instance, the teacher could ask the students to first preview the spellings and meanings of the words before class. During the class, we can invite students to create a story based on the new words.

Regarding listening and speaking, special attention is being paid in our teaching curriculum to these two skills. In the listening and speaking class, listening practices, presentations, group discussions and role plays are included. Students would have many opportunities to practice these two skills.

The assessment for listening and speaking will be a listening test and an impromptu speech conducted by the examiners and students face to face, which will stimulate students to spend more time on these two skills and help them avoid speaking dumb English.

Nevertheless, covert washback still exerts influence because most students are eager to take the CET-4 and they ask a lot about the preparation of the CET-4 and the test-taking skills. In this case, we still need to guide students to practice skills, like skimming and scanning in reading comprehension, and grammar knowledge like word formations, to memorize or distinguish new words. In the meantime, some blank filling exercises and multiple choice questions related to the CET-4 are done in class. Just as Alderson and Wall (1993: 120-121) put in the washback hypothesis, the teaching contents and methods will be affected by a test. This is the case when it is close to the time of the CET-4 exam. Before the CET, teachers would squeeze time to do CET-4 training such as blank filling, multiple choice, translation, and apply exam-oriented methods, sacrificing the time of group discussion and critical thinking. As mentioned before, Alderson and Hamp-Lyon (1996: 281) claim that negative washback will be created when emphasis on skills that require complex thinking or problem-solving is reduced.

As for the students, the majority of them attach crucial importance to the CET-4. As it is mentioned in the introduction, some of them do model tests of CET-4 and memorize CET-4 vocabularies in class, ignoring what is taught in class. This could hinder their English learning as they only learn CET-4 related knowledge. However, to look on the bright side, some of the students complete the CET-4 writing every week and ask their teachers for help. In this case, it would be beneficial for them because the test drives them to persist in their English study. For the English Association in HITSZ, they even organize weekly model tests before the CET-4 to help their fellow students get some training, which can foster the creation of a good learning atmosphere of English study.

All in all, the washback effect of the CET-4, in HITSZ, is generally positive. As for teachers, even though they are under the pressure of the school to provide CET-4 training, they still have the freedom to design their own class. For most of the students, the test drives them to learn English in college. However, for quite a few, the eagerness of success, in the test, blinds their minds and it could impede their comprehensive English study. Therefore, our teachers should guide the students to have a clear sense of the test to maintain the positive washback and reduce the negative impact of the test.

### **Limitations**

The conclusion and implications are based on a literature review of previous studies and local evidence from teaching experience in HITSZ. As objective researchers, the writers know that conclusions cannot be reached without empirical research. At the same time, CET washback is very complicated and changeable within various contexts and different regions. Variations exist with different teachers and students. As Alderson and Hamp-Lyons (1996) mentioned above, a “test will have different amounts and types of washback on some teachers and learners than on other teachers and learners”. (296)

### **Conclusion and Implication**

In this paper, the CET washback has been explored in terms of definitions, classifications and the theoretical backgrounds. Scope, pattern and models have been put forward by scholars and researchers in the field of language testing. Empirical studies were also carried out internationally and domestically. In the large picture, the CETs are set to test students' English level to see their fulfillment of NCETS. However, the educational system, society, and policy makers sometimes misuse this test, leading to a series of problems related to negative washback. Implications to promote positive washback and the decline of negative washback would be:

1. From the perspective of government authority and policy makers, relevant departments should consider the development of new testing systems to adapt to the social needs of personnel assessment, so that the CET-4 would not be a high-stake instrument and be misused by other agencies. College English assessment systems should be encouraged to develop exams for respective schools and reduce the external pressure linked to the student's degree and on teaching. It is necessary to further study the effects of different content and formats of tests on the students in different regions. The CET-SET should be promoted to put emphasis on students' oral ability. (Wang, 2008 )
2. For teachers, we should experiment with effective teaching methods and content to guide students to establish correct learning objectives, correct attitudes toward learning, and the CET. Teachers should cultivate and develop students' intrinsic learning motivation, improve their interests in learning, and develop their learning methods.

3. As for students, they should pay less attention to the CET-4 as an end result and try to develop the intrinsic motivation to develop their learning. Language skills should be promoted in a more inclusive way, with special focus on listening and speaking.

In conclusion, after the reform in 2005, situations have changed. Most research done in China suggests that the positive washback outweighs the negative one. Benefits like promoting college English teaching and learning dominate the field. Similarly, the same case is found in HITSZ. Though causing some problems with negative washback, the CET-4, as a test to evaluate students' English proficiency, has proved to be valuable for both teaching and learning. Let's hope that we can decrease the negative impact and provide more positive washback of the CET-4 in the college English classroom.

### Funding

The authors received financial support from the Harbin Institute of Technology (Shenzhen) Course Construction Project, reference number HITSZUCP17006.

### References:

- Alderson, J. C., & Hamp-Lyons, L. (1996). TOEFL preparation courses: a study of washback. *Language Testing*, 13(3), 280-297.
- Alderson, J. C., & Wall, D. (1993). Does washback exist?. *Applied Linguistics*, 14(2), 115-129.
- Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice: designing and developing useful language tests*. Oxford University Press.
- Bailey, K. M. (1996). Working for washback: a review of the washback concept in language testing. *Language Testing*, 13(3), 257-279.
- Buck, G. (1988). Testing listening comprehension in Japanese university entrance examinations. *JALT Journal*, 10, 15-42.
- Cheng, L. (1997). How does washback influence teaching? Implications for Hong Kong. *Language & Education*, 11(1), 38-54.
- Cheng, L. (2007). *Washback, Impact and Consequences*. *Encyclopedia of Language and Education (2nd ed.): Vol. 7. Language testing and assessment*.
- Department of Higher Education, MOE (Ministry of Education). (2007). *College English curriculum requirements*. Shanghai: Shanghai Foreign Language Education Press.
- Frederiksen, J. R., & Collins, A. (1989). A systems approach to educational testing. *Educational Researcher*, 18(9), 27-32.
- Gu. (2005). *Positive or Negative? An Empirical Study of CET Washback on College English Teaching and Learning in China*. (Doctoral dissertation, Shanghai Jiao Tong University).
- Gu & Peng. (2010). A diachronic study of College English teachers' understanding of CET and its washback effect. *Foreign Languages and Their Teaching* (6), 37-41.
- Gu, Yang & Liu. (2013). A study on the Washback Effect of CET on College English Classroom Teaching- revisits three College English teachers' classrooms. *Foreign Language Testing and Teaching* (1), 18-29.
- Gu, Zhang & Liu. (2014). An Empirical Study of the Innovated CET Washback on Students' Extra-curricular Learning Processes: Based on Students' Learning Diaries. *Journal of PLA University of Foreign Languages (Social Science Edition)* 37(5), 32-39.
- Heaton, J. B., & Han, B.C. (2012). *Writing English language tests*. Beijing: Foreign Language Teaching and Research Press.
- Huang, Xiaoping & Hui, Zhongfeng. (2015). "On the Validity of Educational Evaluation and its Construction" *Higher Education Studies*. 5(4).99-100.
- Hughes, A. (1989). *Testing for Language Teachers*. Cambridge University Press.
- Hughes, A. (1993). *Backwash and TOEFL 2000*. Unpublished manuscript, University of Reading.
- Jin. (2004). CET 4 and 6: Reform in Progress. *Foreign Languages in China*. (1), 27-29.
- Jin, Y. (2005). Giving full play to the positive role of CET 4&6 -- an analysis of the reform plan of the CET 4&6 (for Trial Implementation). *China Higher Education* (9)33-34.
- Li. (2011). *Washback Effect of CET-4 and College English Teaching*. (Master dissertation, Shandong Normal University).

- Liu and Gu. (2013). A review of the international and national empirical researches on washback in the past 20 years. *Foreign Language Testing and Teaching*(1), 6-19.
- Messick, S. (1996). Validity and washback in language testing. *Language Testing*, 13(3), 241-256.
- MOE (Ministry of Education). (2005). The Reform Plan of CET 4 and 6 (for Trial Implementation).
- Niu, Chen & Li. (2011). The Washback Effect of CET-4 of 710 Points System on Listening Teaching based on 4P Model. *Journal of Changchun Normal University(Humanities and Social Sciences)*30(9), 157-161.
- Prodromou, L. (1995). The backwash effect: from testing to teaching. *ELT Journal*, 49(1), 13-25.
- Qi, L.X. (2004). *The Intended Washback Effect of the National Matriculation English Test in China: Intentions and Reality*. Beijing: Foreign Language Teaching and Research Press.
- Shohamy, E. (1992). Beyond proficiency testing: a diagnostic feedback testing model for assessing foreign language learning. *The Modern Language Journal*, 76(4), 513-521.
- Shohamy, E., & Others, A. (1996). Test impact revisited: washback effect over time. *Language Testing*, 13(3), 298-317.
- Vernon, P. E. (1962). The measurement of abilities. *Academic Medicine*,6(37), 159.
- Wall, D., & Alderson, J. C. (1993). Examining washback: the Sri Lankan impact study. *Language Testing*, 10(1), 41-69.
- Wang. (2008). A Systems Approach to the Reform of College English Testing— Report on the “Survey of College English Testing Reform”. *Foreign Languages in China*. 5(4), 4-12.
- Watanabe, Y. (1996). Does grammar translation come from the entrance examination? Preliminary findings from classroom-based research. *Language Testing*, 13(3), 318-333.
- Watanabe, Y. (2000). Washback effects of the English section of Japanese University Entrance Examinations on instructions in pre-college level EFL. *Language Testing Update*, 27, 42-47.
- Watanabe, Y. (2001). Does the university entrance examination motivate learners?: A case study of learner interviews. In Faculty of Education and Human Studies Akita University (ed.) *Trans-Equator Exchanges a Collection of Academic Papers in Honor of Professor David E. Ingram* (pp100-109). Akita University, Japan.
- Xie. (2007). An Analytical Study of Problems with CET-4 from the Perspective of Washback Effect. (Master dissertation, Harbin Institute of Technology).
- Yang. (2003). The 15 years of the CET and Its Impact on Teaching. *Journal of Foreign Languages* (3), 21-29.
- Yang & Tan. (2007). A survey on Washback Effect of CET-4 Writing Assessment. *Journal of Huazhong Agricultural University (Social Science Edition)* (6), 154-157.
- Yu, H.R. (1999). What kinds of personnel of foreign languages are needed in the 21<sup>st</sup> century? An issue beginning with the negative effect of CET-4 and CET-6. *Journal of Jinling Polytechnic College (Social Science Edition)* (4), 33-36.
- Zhou & Liu. (2011). The impact of the use of foreign language test results on test participants-- a case study of CET 4. *Foreign Language Testing and Teaching* (2), 18-29.
- Zhu, Jiangang. (2016). "Higher Education as a Common Good in China: A Case Study for Ideas and Practice" *Higher Learning Research Communications* 6(2), 3-4.