

Successful National English Proficiency Testing Systems in Mainland China and Taiwan

Paper presented at IAEA 35th Annual Conference (9/2009) in Brisbane Australia

by Dr. Byron Gong

byrongong@yahoo.com

Soochow University, Taipei, Taiwan

Abstract

While national college English proficiency tests are being carried out in virtually all universities in Mainland China and some universities in Taiwan, test designers and college English teachers are rising to the challenge of the enhanced responsibilities in terms of the backwash effects of the standardized English proficiency tests, and the accountability of such a testing system. This paper presents a comparative study of the nationwide standardized English language tests at universities in China and Taiwan by introducing how college English education is assessed in Mainland China and Taiwan respectively. To what extent that a large-scale test is designed to measure the intended test contents effectively and satisfactorily can be considered the primary concern for language test designers. In this sense, Chinese universities across the Taiwan Strait have developed different approaches to successful evaluation of their college students' English language proficiency. These approaches, varied as they are in focus and emphasis, are generally linked with educational policy-making and needs of test users at different levels.

This paper analyzes current practices of college ELT evaluation, and offers valuable suggestions. It is believed that a unified English language testing system should be well designed, and such a standardized testing system would bring about much beneficial backwash effects on the ELT programmes at most Chinese universities. The significance of this paper is to raise the issue that most Chinese universities, and those in Taiwan, need without delay a unified English language testing system for quality college ELT results.

I. Introduction

The question of whether a standardized English language testing system is needed for the evaluation of college students' English proficiency is an issue of immense importance for test policy makers and particularly for test users. No matter whether a standardized testing system is conducted at national or collegiate level, such standardized tests may turn out to be a quite high-stakes test for test takers at the grass-roots level, and also important for all test users. Therefore, it is commonly accepted that high stakes language testing should provide as much as possible the fair, reliable, valid, and accountable test results that should have much beneficial backwash effect on both the teaching practice and learning performance, and even on social mobility. However, this could be the ultimate goal for many test designers to achieve in practice.

The difficulty in establishing a sound foreign language testing system at tertiary level can be considered from different aspects and at different levels. Regarding a high stakes foreign language testing system, some countries have organized a national testing team that is responsible for designing and carrying out a unified national standardized English language test periodically, while the educational authorities in other countries or areas may leave this testing task to each individual university to carry out the evaluation of students' English language proficiency.

Both approaches may have their own advantages and may serve different purposes; however, as far as national language proficiency tests to be delivered by thousands of colleges and universities in Mainland China are concerned, a test conducted by each university providing its incomparable test scores with those of other universities could have more problems due to the lack of a unified testing criterion of test validity when compared with other universities at national level. The situation could be even more problematic regarding a graduation test of the

English language proficiency in a university if each basic teaching unit conducted its own language testing with different criteria. On the other hand, socio-economically, some test users such as employers and government organizations may also feel perplexed by incomparable test scores provided by different colleges and universities, not to say the credibility of some universities is far from being accountable. Without a valid or an accountable testing system, real hard-working college graduates would become the sacrifice for such an individualized testing system. In this respect, universities in Mainland China and Taiwan have come across this problem for long, and many of them have to work out their own methods to cope with this issue unless some universities do not consider it necessary that a national testing system should be established.

The main purpose of this paper is to introduce how college English education is assessed in Mainland China and Taiwan respectively. To what extent that a large-scale test is designed to measure the intended test contents effectively and satisfactorily can be considered the primary concern for language test designers. In this sense, Chinese universities across the Taiwan Strait have developed different approaches to successful evaluation of their college students' English language proficiency. These approaches, varied as they are in focus and emphasis, are generally linked with educational policy-making, needs and social weight at different levels.

II. Theoretical Considerations

The theoretical framework of this paper is based on the conception that a successful language testing system should not be viewed only from the aspects of test reliability and validity, but also test usefulness or accountability in a broader socio-economic context. Fundamentally, this is because national education is to serve the main interest of the people, i.e. represented as the national interest, rather than the interest of a special group, either personal academic interest or private research interest.

By usefulness, the writer means that a large-scale or national test should serve the interests of test users, and the test construct should be linked with the needs of socio-economic needs at both the macro level and at the micro level. Test usefulness could be the most important consideration in designing and developing a language test (Bachman and Palmer, 1996:18). In other words, test usefulness is to be linked with a specific purpose, a particular group of test takers and a specific language use domain in mind. In this sense, a language testing system can be considered successful only when its main users, or stakeholders, accept it as useful and accountable. Yang (2009) also points out that; "Test results are to be used by the public; from the sociological point of view the social weight of a language test is determined by the popularity of its test users."

Practically, take universities in both Mainland China and Taiwan for example, whether a unified college English testing system should be adopted depends on how useful such a language testing system can serve the interest of its users. However, no matter what kind of college English testing system is to be adopted, these Chinese universities need to be accountable for the English language testing system they use.

Then, the writer would raise a fundamental question: Is it more advantageous for universities in Mainland China and Taiwan to use a kind of unified and standardized English language testing system as a threshold graduation test? The tentative answer is positive. A hypothesis in this study holds that a standardized English testing system can enhance not only the test quality regarding both reliability and validity, but also can consolidate the quality college ELT programmes in terms of test constructs and usefulness for the state interest as well, not to say the huge positive washback effects on national college ELT programmes. The fact that a unified and standardized college testing system has been smoothly and intensely carried out in different provinces in Mainland China for over 20 years could be a good case for college ELT programs in Taiwan and other areas to draw lessons. Next, we shall take a look at how college graduation English tests are conducted in Mainland China and Taiwan respectively.

III. Graduation English Tests at Universities in Mainland China

It is important to note that all colleges and universities in Mainland China today are public and they are under the leadership of educational authorities at both national and provincial (or local) levels. On the other hand, we also need to know there has been a considerable change in China's tertiary education since China carried out its policy of economic reform and opening to the outside world 20 years ago. Today, most of the educational authorities are a group of people who have received modern higher education with a PhD degree from either at home or abroad. Such a situation is significantly different from that of 20 years ago. Currently, college ELT in Mainland China can be described as both flourishing and problematic in terms of its educational reform, evaluation, funds for college development, teacher training, and teaching material. Meanwhile, colleges and universities have been actively engaged in various international cooperation and exchange programmes.

Foreign language education, virtually English language education, has been intensely carried out at all levels of education in China. English is a required subject in the national entrance examinations of higher education for all types of colleges and universities. English is also a compulsory course for all non-English majors to take for at least two years. Meanwhile, the traditional conception of examinations for social mobility is widely accepted and highly respected in the Chinese society.

Regarding China's educational system and its relevant examinations, we can hardly be wrong by saying that China is a place full of various examinations. In view of large-scale English examinations in China, there are two types in general. One type of English examinations is that performed by foreign testing organizations. The typical examinations of this category in China are IELTS, TOEFL, GRE, GMAT, Cambridge Business Examinations, BULATS, etc. The other type of English examinations is delivered by China's Ministry of Education and its subordinating organization—the National Education Examinations Authority (NEEA). In addition, there are numerous national English examinations (including those for ESP and vocational) designed for different purposes by different government ministries and educational authorities at national level, not to mention hundreds of those English examinations at provincial or local universities. However, as for higher education in general, there are various problems among most of these English tests at national level. China's National Education Examinations Authority realized such a chaotic situation and admitted as early as 20 years ago:

“Currently, the National Education Examinations Authority (NEEA) administers a wide variety of different English examinations with a total of well over 30 papers, more than ten of which are public English examinations; these were individually designed for special purposes by authorities in charge as a result of different demands under special historical conditions after the Cultural Revolution. The exams are not placed on a common scale and there is a lack of consistency in standards with different standards being used at the same grade and different levels of testing sometimes being used with the same standard.” (NEEA 2001)

The above quotation shows that the Chinese examination authorities realized that the situation of English language testing services in China 20 years ago could hardly meet educational and social demands. Then, in the light of such a situation, there is a call for a kind of unified and standardized national English language testing system for colleges and universities to adopt so that college students need to take as a threshold English proficiency test for graduation.

More college teachers (Yang and Weir, 1998) believe that without a unified and standardized testing system, the quality of nationwide college ELT programs could hardly be achieved or guaranteed. Consequently, college teachers feel that setting up a nationwide college English testing system is necessary, and then educational authorities have risen to the challenge of their responsibilities.

With this background, the NEEA, with the help from the British Council, launched a new national college English proficiency testing system in 1987, and this college English test (CET) is comprised of four parts: 1) Listening Comprehension, 2) Reading Comprehension, 3) Cloze, 4) Writing and Translation. The college English test (CET) has been carried out successfully in all

provinces in Mainland China for over 20 years, and has been highly recommended by the Ministry of Education in Mainland China although it is not a prerequisite for a bachelor's degree. Today, among all the various kinds of English proficiency tests, the CET has become such a significant test that its test results are highly accepted by the public and society during the past decade. Employers in China prefer applicants with a CET certificate. As Liying Cheng pointed out (2008):

“To obtain a bachelor's degree in Chinese universities, these students often need to pass the College English Test – an English language proficiency test. English is an examination subject for all students who wish to pursue a graduate degree in China.... In reality, being successful in the various English tests and examinations is the key to the success in life for many in China.”

It is no exaggeration to say that the educational authority in Mainland China has been promoting an unprecedented English language testing system at universities with millions of test candidates each year. The national standardized CET test, which was revised in 2005 and 2007, has been such a high-stakes college English proficiency test that millions of undergraduate university students in China are required to take before their graduation (Han et al., 2004). The number of CET candidates is on the increase every year. In the 1995 academic year, 583,135 students in China took the CET, with a passing rate of 66% (Yang and Weir, 1998); and 9.58 million students took the test in the 2005 academic year (Jin, 2005). Considering the huge number of CET candidates, it is clear that China has not only the largest English test candidates, but also the largest English-learning population in the world. Although China's Ministry of Education altered its test policy in 2005 by stating that the CET is not to be directly related to college students' graduation, college students still consider the CET test crucial because the CET test certificate is an important criterion for many employers to consider at a job interview. In other words, practically, it is an irreversible trend for millions of Chinese college students to take the CET test before their graduation. Although there are negative voices against the CET (Han et al., 2004), it is generally believed that such a standardized English language testing system has brought about a cumulative positive effect on the quality teaching of college English education in Mainland China (Yang and Jin 2009; Jin, 2005; Yang and Weir, 1998). Therefore, as being linked with both educational and social status, the CET test has become high-stakes for 20 years since its first launch in 1987. (Notice: English-majors have a more advanced national English proficiency test, i.e. TEM-8, the Test for English Majors, to take at the end of the 4th-year in their undergraduate programme.)

CET Accountability

As the writer states in his theoretical framework, a successful language testing system should not be viewed narrowly from the aspects of test reliability and validity only with some statistic figures, but also viewed from test usefulness or accountability in a broader socio-economic context. As a national test among all the universities in Mainland China, the CET is accountable to the public with the following five features: (http://www.en.cet.edu.cn/overv_displaynews.asp?id=277):

1) The CET has been widely accepted and supported among universities in Mainland China.

The CET has received much attention both from institutions of higher learning at all levels and from relevant educational departments and they have also stimulated and motivated teachers and students in the process of English teaching and learning. Large amounts of statistical data and experimental evidence collected through research have proved CET that has not only high credibility, but also great effectiveness. They have met the quality standards required by large-scale standard examinations and, based on the requirement of the syllabus, could well reflect the English proficiency of the college students in China. Therefore they have greatly promoted the implementation of college English teaching program and improved the teaching of college English as well.

- 2) Statistically, the CET has got solid data to prove itself as reliable and valid. The CET presents an annual objective data description of the English proficiency of college students in China. The modern educational statistical approaches have been adopted extensively in the CET system, in which test scores are processed by means of equivalence method annually.
- 3) The CET provides important reference data for universities. The CET testing system provides tremendous information after each examination, which has provided a dynamic basis for the educational administrative departments at all levels to make sound decisions and also offers feedback information for various institutions of higher learning to take appropriate measures to improve their teaching quality accordingly.
- 4) The CET has got high academic and public reputation. A set of complete procedures have been formed in terms of question setting and examining, organization of the examination, statistical analysis and release of scores of the CET. Thus the CET belongs to one of the large-scale examinations that are well-organized and conducted systematically according to the quality requirements of standard examinations.
- 5) The CET is socially well accepted in Mainland China. Widely accepted by society, the CET has served as one of the preconditions for the personnel departments at various levels to take on college graduates. In this way they have produced certain social benefits.

Construct Validation for Successful CET

According to the CET Committee, it is claimed that the CET is designed both criterion-related and norm-referenced (Yang and Weir, 1998). Its test specifications are supposed to reflect the requirements set in the National English Teaching Syllabus, and its test results should provide a valid score ranking of test candidates for all the stakeholders. As Alderson pointed out (1995), it is generally regarded as the correct method by analyzing test specifications as a starting point in studying the construct validity of a test. Hence, it is also important to take a look at how the CET can be considered successful regarding its construct validation. To what degree that the test specifications of the CET can reflect the intended requirements of China's national English syllabus of higher education is considerably relevant to the degree of how successful or valuable the CET could be. In this aspect, the CET criteria are based on China's National College English Teaching Syllabus for Non-English Majors 1999 (which was revised and called "Requirements" in 2007). This national syllabus stipulates specific quantitative requirements for college students to achieve in terms of their English language proficiency, and skills of reading and listening are of paramount importance. (<http://edu.people.com.cn/GB/8216/43375/5995154.html>).

The CET tests have two basic versions, CET-4, and CET-6. The CET-6 is for students who have passed the CET-4, and have taken elective English course of Band 5-6. The CET Spoken English Test (CET-SET) is administered only to a very small number of students who want to take by themselves on the condition that these students have passed the CET-4 with a score of 80 or above out of a full score of 100, or the CET-6 with a score of 75 or above. However, only the CET-4, which is the focus of this paper, is considered as the benchmark test that virtually all undergraduate students need to pass, and the CET test is administered twice a year, in January and June. According to China's National College English Testing Committee (2006), there are four main components in the CET-4 test: Listening Comprehension (35% - short dialogues and long authentic talks), Reading Comprehension (35% - careful reading and fast reading), Cloze (15% - one cloze and sentence translation), and Writing (15% - one short essay).

As for test specifications of the CET, the study report states that the guiding principle is to reflect the requirements of the national English teaching syllabus. According to the research report by Jin Yan (2006) and Yang and Weir (1998), the CET test specifications have very well met the requirements of the national English syllabus. As the test is to help to implement this national syllabus, the CET test designers had paid much attention to the following aspects so as to construct a theoretical framework for the CET test:

- 1) The relationship between knowledge and ability: This means, conceptually, language is a tool for communication. The ultimate aim of EFT is to ensure that students can use English to communicate. Therefore, the CET should test more language skills rather than language

- knowledge.
- 2) The relationship between fluency and accuracy: The designers of the CET have set specific speed requirements of reading, listening and writing (i.e. 50 wpm and 129 wpm are set for reading and listening in the CET-4 test).
 - 3) The relationship between sentence understanding and discourse comprehension: As communication is based on discourse comprehension, the CET should not only take into consideration of sentence structures, but also the ability to understand discourse.
 - 4) The relationship between receptive ability and productive ability: This means the CET specifications require that both passive and active skills are to be examined.

According to the research by Yang and Jin (2009), Jin Yan (2005) and Yang and Weir (1998), the CET test is designed according to the above four major considerations which constitute the basis for its construct validation at a macro level. The construct validation of the CET test is therefore can be indicated by the degree of how the CET test can be accepted by both the educational authorities and university English teachers. According to an official survey by China's National College English Testing Committee (2006), the CET has successfully achieved the aims of its test specifications; and the construct validation based on the theoretical framework can be well represented in each delivery of the CET test. Specifically, the statistics also provide the following implications:

- The internal reliability of objective items in the CET test reaches 0.9 or above every time when the CET test is conducted, indicating that the reliability of the CET is high.
- A series of studies of questionnaires on the CET has indicated that 92% of college teachers in China agree that the CET test can effectively reflect students' actual English proficiency level, indicating a high validity in terms of expert judgment.
- As the CET is a criterion-related norm-referenced test, the passing score set in the CET correlates with the teachers' assessment of the test candidates' passing score with a correlation coefficient of 0.82 ($p < 0.01$). In addition, the CET test scores correlate with the order of class assessment results given by the teachers with a correlation coefficient of 0.7, which is very good because it is difficult to achieve such a high coefficient in large-scale standardized tests.
- Over 86% of college teachers agree that the contents of the CET are appropriately designed and each part has a proper weighting.
- The CET has a complete testing system, including item bank management, test formation and organization, administration, statistical analysis of test results, test fairness, and practicality.

Therefore, upon conclusion of this part, the writer maintains that the construct validity of the CET is mainly associated with the state interest at national level. As Ross (2008) points out, the CET is designed to provide a dual function of criterion-related and norm-referenced standard for assessing English language proficiency. It is accountable why the CET can be considered successful regarding the fact that the national English syllabus is adequately reflected in the CET test constructs. On the other hand, the feature of being norm-referenced enables the CET to provide an accountable and valid ranking of the test candidates, which has provided a valuable information for employers across mainland China. In this respect, the CET's socio-economic role can not be replaced by any other English language testing system so far. Therefore, the writer of this paper maintains that although there have been ups and downs, the CET has proved itself a very successful and valid college English proficiency test in Mainland China during the past two decades.

Next, the college graduation benchmark testing system in Taiwan will be discussed.

IV. Graduation English Tests at Universities in Taiwan

Regarding the college English graduation test, there is a very different scenario for college students and university authorities in Taiwan. Luckily, there is not a mandatory island-wide English proficiency test set by Taiwan's Ministry of Education for undergraduate students to take

for graduation. Unlike its counterpart in the Mainland, the Chinese educational authority in Taiwan has been carrying out an American style of autonomous and decentralized language assessment within colleges and universities. There is no national English syllabus set by Taiwan's Ministry of Education (MOE) for colleges and universities to follow. Instead, the MOE has transferred power to lower levels at each individual university, which gives more freedom to universities to decide what kind of English proficiency is needed for their undergraduate students according to each university's own principles; and many universities in Taiwan have recently announced that they will carry out their own benchmark English testing system for college students to take. In addition, undergraduate students can also take external English proficiency tests as a proof of their English proficiency before they finish the 4-year university education, such as the GEPT test (General English Proficiency Test, a criterion-referenced test set by a non-governmental organization in Taiwan), or TOEIC, IELTS, and TOEFL. Nevertheless, although there is the CSEPT (College Student English Proficiency Test, which is new), the local GEPT test is the most popular English proficiency test for college students to take (LTTC, 2007).

Meanwhile, among the English proficiency tests designed by individual universities at grass-roots level in Taiwan, different universities have their own testing systems and criteria. The SCUEPT test (Soochow University English Proficiency Test) appears to be at the forefront of the campaign for a standardized English proficiency threshold test for undergraduate students to take for graduation. Many other universities are also trying to design their own proficiency tests now. By and large, it is clear that a variety of benchmark English tests will soon become high-stakes tests for thousands of college students in Taiwan to take. Socio-economically, such a test certificate would help college graduates to have better opportunities in the job market, too.

Construct Validation in the GEPT/SCUEPT Test Specifications (Taiwan)

Now let us take a look at the benchmark English test for college students in Taiwan. There are more than 150 officially accredited universities in Taiwan. However, there lacks a cohesive paradigm of college English assessment at the tertiary level. As mentioned earlier, there exists no requirement from the educational authorities in Taiwan to demand all college graduates to take an English proficiency test before they graduate. According to a 2007-year report on the GEPT, only 22% of GEPT testees took the GEPT in order to give their test scores to their universities for reference. As for the specifications, the GEPT is not designed to test just college students' English proficiency, but to test the English proficiency of the general public, which is very different from that of the CET.

Notwithstanding this, the freedom from government control regarding using a standardized and centralized assessment within colleges and universities in Taiwan reflects autonomy in defining the construct of ability, whose rationale could have a different framework for different socioeconomic purposes. Practically, the general situation of the evaluation of English programmes among universities in Taiwan is that test scores may be inconsistent and incomparable. Different universities adopt their own methods in evaluating students' English proficiency, and such methods may be inconsistent each year, and even differ from one individual teacher to another. Therefore, the evaluation results conducted by different universities are difficult to interpret in terms of statistical analysis at national level.

At present, graduation examinations across universities are mostly of progress tests or achievement tests, and the contents of such achievement tests could be widely different from one university to another (Gong 2004). For example, let us look at the achievement test scores of the same English course at two campuses of the same one university.

Table 1: Pearson Correlation of Freshmen English between Taipei and Kaohsiung Campuses (N=900; 1/2005)

	Taipei Campus	Kaohsiung Campus
Taipei Campus	1.000	0.034
Kaohsiung Campus	0.034	1.000

p<0.01

Table 1 reflects that the test scores are not comparable even within the same university. No

significant correlation can be found between the test scores at its two campuses (Kaohsiung and Taipei) of the same university, with a Pearson correlation of 0.034 ($p < 0.01$). This may suggest that as there are no national English syllabus and test specifications for universities in Taiwan to follow, college teachers can use different test formats and constructs for evaluation at will.

With an autonomous and decentralized educational system, each individual college or university may decide their own testing criteria according to their own needed rationale. But for universities at grass-roots level, their test results appear to be incomparable. Thus, the English language testing is said satisfactory and successful only in terms of the interpretation of the needed test construct, or “ability” by each university itself.

VI Discussion of Empirical Data of the CET, GEPT, and SCUEPT

Test designers in both Mainland China and Taiwan paid much attention to the issue of validity in their tests at different levels. Let us look at some empirical data so as to have a better view of the CET-4 test and GEPT (or SCUEPT). First, reliability results of the above three tests are reported as follows: the reliability (Cronbach's alpha) of the CET is 0.9, and 0.86 for the SCUEPT, while the GEPT is said to be 0.85 (LTTC 2003:23), fairly good. Furthermore, the internal correlation coefficients of these tests could be very useful for us to evaluate construct validation of these tests. The following Tables 2-5 show the internal correlation coefficients.

Table 2: Internal Correlation Coefficients of the CET-4 (Yang and Weir, 1998:60)

	LC	RC	VS	CL	WR	Total
Listening	1.000					
Reading	0.563**	1.000				
Vocabulary	0.539	0.615	1.000			
Cloze	0.467	0.531	0.626	1.000		
Writing	0.388	0.359	0.470	0.404	1.000	
Total	0.792	0.892	0.802	0.707	0.581	1.000

** : Correlation is significant at the 0.01 level (2-tailed).

Table 3: Internal Correlation Coefficients of the GEPT (High-Intermediate Report, 2007, p.2)

	Listening	Reading	Writing	(Speaking)
Listening	1.00			
Reading	0.37**	1.00		
Writing	0.19	0.32	1.00	
Speaking	0.39	0.23	0.38	1.00

$P < 0.05$

The above Table 2 shows that the internal correlation coefficients of each part in the CET-4 are between 0.3-0.7, which is fairly good. Compared with Table 2, Table 3 shows that the internal correlation coefficients in GEPT in Taiwan is slightly less desirable, especially the coefficient between Listening and Writing in GEPT (2007) is on the small side compared with that of the CET shown in Table 2. But as we mentioned earlier that the statistics of the GEPT may not be constant, its earlier 2000 report provided a different analysis than the 2007 report, which can be seen in the following Table 4.

Table 4: Internal Correlation Coefficients of the GEPT (High-Intermediate 2000 LTTC Report, p. R-11)

Sub-test	Reading Part A	Reading Part B	Reading Part C
Reading Part A	1.000		
Reading Part B	0.681	1.000	
Reading Part C	0.686	0.722	1.000
Listening Part A	0.591	0.590	0.598
Listening Part B	0.629	0.624	0.648
Listening Part C	0.620	0.605	0.680

$p \leq 0.01$, $N = 375$

The internal correlation coefficients of the GEPT (2000 test) appear to be more convergent than

that of 2007 test.

The following Table 5 is about the internal correlation coefficients of the SCUEPT test at Soochow University in Taiwan conducted in May 2008.

Table 5: Internal Correlation Coefficients of SCUEPT (Pearson / N=2004, 2008)

	Sentence	Fast reading	Careful reading	Cloze	Short conversation
Fast reading	0.256**				
Careful reading	0.320**	0.327**			
Cloze	0.305**	0.249**	0.362**		
Short conversation	0.384**	0.322**	0.436**	0.360**	
Long talks	0.312**	0.291**	0.426**	0.283**	0.611**

** : Correlation is significant at the 0.01 level (2-tailed).

Looking at Tables 2 and 5, we can see the inter-subtest correlations of the GEPT and SCUEPT in Taiwan generally appear to be on the small side when compared with those of the CET. This might indicate that each inter-subtest in GEPT or SCUEPT is too disintegrative regarding language communication skills. Probably, further efforts are needed so as to achieve less divergent construct validity in the GEPT.

VII. Final Discussion

So far we have briefly discussed the features and accountability of the College English Test (CET) in Mainland China and the General English Proficiency Test (GEPT), and SCUEPT (a representative English proficiency test designed by a local university) in Taiwan. We can see the educational authorities in Mainland China and Taiwan have different rationale for the most suitable testing systems for the college ELT programmes in their own socioeconomic contexts. The findings of this study have revealed that positive backwash effects on national ELT programmes can be enhanced effectively when construct validation can be linked with not only a needed rationale but also strong government support and cooperation from different educational and socio-economic forces. As for the research question raised in this paper, i.e. *Is it more advantageous for universities in Mainland China and Taiwan to use a kind of unified and standardized English language testing system as a threshold graduation test?* The writer would hold that national college English testing in Mainland China has played a very positive role in her national tertiary English programmes, and the standardized CET is quite successful in terms of both educational and social aspects, which is good reference for their counterparts in Taiwan.

At macro level, for the Chinese educational authorities in the Mainland, the CET test serves the state interest and the needs of the national higher education. The CET reflects governmental initiatives for centralization and standardization of language testing at a national level, with a centralized definition of ability construct. The CET appears to be such a mandatory standardized English language testing system that virtually all college students in the Mainland need to pass this high-stakes benchmark test for graduation. Therefore, the CET has greatly mobilized all Chinese college students to study English hard. Although there are criticisms of the CET (Han et al., 2004), the major positive backwash effects of the CET test is that all universities in China have realized the importance of college English education, and have taken various actions to promote college English actively, which has brought about a cumulative positive effect on the quality teaching of college English education nationwide, not to say the heavy social influence of the CET (Jin and Yang, 2006). The state English language testing policy in China, generally, exerts a more desirable influence on her national college English programmes. Its CET testing system has produced far-reaching positive effects on college ELT programmes in turn. In this

sense, the CET is quite successful due to its immense function of meeting the needs of the national English teaching syllabus and getting college students mobilized to study English.

Meanwhile, the successful feature of college English testing in Taiwan can be reflected from the perspective that assessment is more linked with local conditions. The universities in Taiwan have much more freedom to decide the criteria of what English proficiency is needed for their undergraduate students according to each university's own understanding and interpretation of test construct abilities. Many college teachers and researchers in Taiwan do not want to see that testing drives teaching. Therefore, the decentralized educational policy in Taiwan has successfully guaranteed freedom of keeping autonomous testing policy at each individual university, which has greatly motivated the initiative to design an English test mostly suitable to the grass-roots needs and conditions. On the credit side, college teachers and researchers in Taiwan are also making greater efforts to understand the relationship between teaching and testing, but the approach they take differs significantly from that of the Mainland China.

Conclusion

As a final point, a standardized benchmark college English test might provide an efficient and valid solution to the malaise of the English language tests at universities in Mainland China. However, the decentralized approach to college English assessment in Taiwan could be especially meaningful when construct validation is considered at grass-roots college level, rather than at national level for the state interest. Nevertheless, the writer would like to point out that to establish a unified language testing system for the college ELT programmes does not mean that the purpose of college ELT should be hence further linked with students' ability of linguistic code reading for the main purpose of test taking. It is more important for university students to become competent English users than just being able to pass an English test. In this sense, both the CET in Mainland China and GEPT in Taiwan need to improve their validity, especially the ability to use English for communication. As the late Chinese leader Mr. Deng Xiaoping said: no matter whether it is a white cat or a black cat, it is a good cat as long as it can catch mice.

References

- Alderson, J. C., Clapham, C., and Wall, D. (1995) *Language Test Construction and Evaluation*. Cambridge: CUP.
- Bachman, L.F. and Palmer, A.S. (1996) *Language testing in practice*. Oxford: OUP.
- Cheng Liying (2008) The key to success: English language testing in China. *Language Testing* 25(1): 15-37.
- Gong, Byron (2004) "A Need for a Unified Assessment of College English Language Programs—Some Theoretical and Practical Considerations for Quality ELT in Taiwan" *Shih Chien Management Commentary*, Issue 1.
- Jin, Y. and Yang, H. (2006) The English Proficiency of College and University Students in China: As Reflected in the CET in Language, *Culture and Curriculum*, 21–36.
- LTTC (2007, 2003, 2000) *A Statistical Report on the Scores of a GEPT® Test*. Taipei, LTTC.
- National College English Testing Committee, PRC (2006) *College English Test Sample Papers*. Shanghai, China: Shanghai Foreign Language Education Press.
- Ross, Steven J. (2008) Language testing in Asia: Evolution, innovation, and policy challenges. *Language Testing* 2008 25 (1) 5-3.
- Yang, H. and Weir, C. (1998) *The Validation Study of the National College English Test*, third edition. Shanghai, China: Shanghai Foreign Language Education Press.