On Washback of Testing to General English Education

Tang Sheng-ping & Xu Chong-ning

Chongqing Technology & Business University

Abstract:
Language testing, as a means of assessment system, is an important part of ELT (English Learning and Teaching). The educational society identifies the influence of important tests on ELT and refers to it as washback. Although there is controversy about its beneficial or harmful effects on teaching, it is generally accepted that testing generates both negative and positive washback. By making a comparison between and an analysis of the washback from the three national English tests (NMET, CET-4, CET-6) in the general English education, this paper deals with how to take advantage of the positive washback and minimize the negative one to promote ELT.

Key words: testing; English teaching; washback

The educational society identifies that important tests have on teaching and learning great effects which are referred to as washback, a term popular in British applied linguistics and commonly termed as backwash in the field of general education. The concept of washback presupposes a belief in the notion that tests are prominent determiners of classroom practices and events. Alderson & Wall, in making systematic study of washback, put forward 15 washback hypotheses, including the influence of a test on the course characteristics, instructional time, people involved (test-takers, teachers and administrators), the teaching speed, degree and depth, etc. Although there is controversy about its beneficial or harmful effect on teaching and learning, it is generally accepted that testing yields both negative and positive washback. By making a comparison between NMET (the National Matriculation English Tests) and CET-4 / CET-6 (College English Test Band 4 / Band 6) in the general English education and an analysis of their respective washback, this paper deals with how to take advantage of the positive washback and minimize the negative one to promote ELT.

1. A perspective of testing

Language testing, as a measure of evaluation system, is an important part of ELT. The Syllabus of College English in force holds that testing is an effective means of checking its implementation and evaluating teaching quality. It also functions as the main sources for both teachers and students to get feedback, which enables them to reflect on their teaching and learning activities and thus help improve the jobs of both sides. As we know, a scientific assessment system comprises the formation assessment and the end product one. The former applies to students to mirror their performance in the process of learning, and the development of their affects, interests, attitudes, strategies, etc. while the latter, in most cases, assuming the form of testing, examines students’ comprehensive knowledge and language competence, and also serves as an indicator of the teaching quality of a school.

The functions mentioned above apart, testing can also work as a consolidating instrument, spurring students to review and hence solidify what has been learned. Though pragmatic means such as listening, speaking or writing are effective for consolidation, what can’t be ignored is the
motivation testing triggers, for any coming test or examination is most likely to drive students to regular revision, which proves an essential link for knowledge to sink in. From this standpoint, we can perceive explicitly the relationship between knowledge-learning, revision and testing as that between the purpose, the means to internalize knowledge and the drive for revision.

2. The differences and similarities of NMET and CET-4/CET-6

Testing, according to its varied purposes, can be categorized into the proficiency test, the placement test, the achievement test, the diagnostic test, etc. Based on the interpretation of the test results, testing can be grouped into criterion-referenced test and norm-referenced test (R.Glaser 1963). The former is an approach to absolute assessment, intended to test whether the candidates have reached the required level, while the latter is a tool for relative evaluation, applied to streaming the candidates, and hence, a selective test. Apparently, NMET and CET-4/CET-6 share similarities mainly in that they are all proficiency tests but bear striking differences, for NMET, with the mission of streaming candidates into higher-learning institutions of different kinds, is of a placement test and belongs to the norm-referenced category. As an important test, NMET plays a significant role in guiding and backwashing middle school ELT. On the other hand, CET-4 and CET-6, with the purpose to check whether college students have met the requirements set by the syllabus, belong to the criterion-referenced category. Given that passing CET-4 / CET-6 has been set by many a personnel department as one of the basic requirements of selecting candidates (Yang hui-zhong, 2003), they have influenced the employment of college graduates, and thus naturally become their concerns. As influential tests and a means of assessing teaching, they work as a guidance of and yield washback effects on college ELT.

Table 1 makes a comparison between NMET and CET-4/CET-6 to illustrate the differences and similarities in the testing nature, purpose, contents, the proportion of objective and subjective items, grading, and above all, the way of providing feedback information.

<table>
<thead>
<tr>
<th>test</th>
<th>NMET</th>
<th>CET-4</th>
<th>CET-6</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. nature</td>
<td>National proficiency test; norm-referenced category</td>
<td>National proficiency test; criterion-referenced category</td>
<td>The same as CET-4</td>
</tr>
<tr>
<td>2. purpose</td>
<td>For matriculating students and guiding and backwashing middle school ELT</td>
<td>For checking whether the candidates have met the basic requirements of the syllabus; guiding and backwashing college ELT</td>
<td>For checking whether the candidates have met the advanced requirements of the syllabus; guiding and backwashing college ELT</td>
</tr>
<tr>
<td>3. Testing items</td>
<td>LC, VS, RC, CW, cloze and error-correcting</td>
<td>LC, VS, RC CW, apart, there’re alternatively SA, Tr, CD, cloze, with oral English test for 80-plus scorers.</td>
<td>The same as CET-4 in addition to error-correcting, with spoken English test for 75- plus scorers.</td>
</tr>
<tr>
<td>4. Subjective items</td>
<td>With regular items such as error-correcting in a passage and CW, accounting for 23%</td>
<td>With the regular item CW, and alternative ones such as SA, CD, etc., accounting for 25%</td>
<td>The same as CET-4 besides error-correcting accounting for 25%</td>
</tr>
<tr>
<td>5. Objective items</td>
<td>With multiple choice items in LC, VS, RC, and cloze, accounting for 77%</td>
<td>With multiple choice items in LC, VS, RC, and cloze, accounting for 75%</td>
<td>The same as CET-4</td>
</tr>
</tbody>
</table>
6. Scoring

<table>
<thead>
<tr>
<th>Score Range</th>
<th>Number of Significant Marks</th>
</tr>
</thead>
<tbody>
<tr>
<td>60-64.5</td>
<td>3</td>
</tr>
<tr>
<td>65-74.5</td>
<td>4</td>
</tr>
<tr>
<td>75-84.5</td>
<td>5</td>
</tr>
<tr>
<td>85-94.5</td>
<td>6</td>
</tr>
<tr>
<td>95-100</td>
<td>7</td>
</tr>
</tbody>
</table>

7. Feedback means and the influence

<table>
<thead>
<tr>
<th>Feedback Type</th>
<th>Details</th>
<th>Scoring</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Timely</td>
<td>Media of all sorts provide the test paper and answer soon after the exams, by which students can self-assess their own performance and teachers can reflect on their teaching and readjust their approaches to future work.</td>
<td>Full marks: 100, with 60 as the minimum of significant marks.</td>
<td>The same as that of CET4</td>
</tr>
<tr>
<td>Delayed</td>
<td>Neither students nor teachers can get detailed information about the students' performance. The global score announced months later can aid little in teachers' reflection or students' self-assessment.</td>
<td>The same as that of CET4</td>
<td></td>
</tr>
</tbody>
</table>

(Note: LC- Listening comprehension; RC- reading comprehension; SA- short answer questions; VS-vocabulary & structure; CW- composition writing; Tr.-translation; CD-compound dictation)

The differences stand out not only in the purpose, nature, grading, but in other respects as well. For example, to matriculate candidates, the significance of marks for NMET varies annually with the difficulty coefficient, the scale of enrollment, geographical districts, etc., while CET-4/CET-6 have fixed 60 as the minimum of significant points. To avoid test-oriented teaching, they adopt both regular and alternative testing items such as SA(short answer questions), CD(compound dictation), Tr.(translation), cloze work etc., and the proportion of subjective items keeps growing. What deserves mentioning is the significant-point grading system by CET-4/CET-6, which is intended to advocate colleges to bring students up to a higher level of performance rather than settle for raising the passing rate, and also meant to give guidance to different colleges accordingly. But what catches our attention is the conspicuous differences in feedback means.

3. Discussion on Feedback means and the influence

Bailey(1996) stresses four points in creating the positive washback of testing, i.e., the purpose of language-learning, authenticity of the testing, students' autonomy and self-assessment and the feedback of test results. He also puts forward 8 criteria of judging whether a test can yield any positive washback: Do the test participants know the purpose of the test? Is the feedback clear, detailed and timely? Do the test-takers believe in the validity and credibility? Does the test comply with the principles of sound curriculum design? Is the test based on any definite aim and goal? Does the test follow any theoretical principles? Does the test ever employ any authentic testing tasks? Do the test-takers ever engage themselves in self-assessment? Since items 4 to 7 concern the test-composer, this discussion deals with the other four items.

Interviews with hundreds of NMET and CET-4/ CET-6 candidates found that the former’s understanding of item 1 was limited to the NMET’s purpose of matriculating students while the latter claimed that theirs was to the influence of CET-4/ CET-6 on their academic degree and the future employment. Neither mentioned the tests’ influence on teaching and learning. As to item 2, Please refer to Table 1, Point 7. In the case of item 3, roughly a third of NMET candidates complained about the fluctuation of difficulty coefficient, holding that too big or small a difficulty coefficient would affect the differentiation and the dispersion resulting in misjudgment of candidates’ real proficiency thus they doubted its fairness and credibility while CET-4/ CET-6 interviewees held that the validity
and credibility of a test was an issue involving complicated factors but since the society and employing units accept the test they chose to believe its fairness and credibility. As for item compared with NMET, which renders the test paper with key available to the public right after the test via media so as for the candidates to make self-assessment and for teachers to reflect on the teaching and grasp any message of teaching reform conveyed by the paper, CET-4 and CET-6 offer no such accessibility for either the candidates or teachers until months later the authority concerned provides some global feedback to the school, which, honestly speaking, turns out to be of little value to individual students.

Bailey (1996) once commented on the relationship between self-assessment autonomy and the positive washback effect and quoted the explanation of von Elek that self-assessment enables students to assume greater responsibility for self-evaluation of their linguistic proficiency and academic performance find the deficiency in learning, determine the distance between their actual level and the desired one so that they can be further motivated to strive for a more definite goal.

The Assessment Proposal prescribed by the National Education Commission holds that teachers should attach importance to the feedback effect of evaluation on teaching and learning and that teaching assessment is to render teachers access to related feedback information by which they are to improve their teaching through reflection and adjustment. It is a fact that the National CET Committee provides each institution with a lot of feed-back information in a test report including a score list informing the population that has passed or gained excellence the average score and the standard disparity of each sector in the test paper etc. These data have been utilized by the administrators as an instrument for macro-guidance of College ELT Yang Hui-zhong, 2003. Yet, as subjects of the tests both teachers and students need more detailed individual-related feedback for students to see the difference between their own proficiency and the criterion so that they can timely set the next goal and for teachers to analyze the test results tracing the cause of the existing problems so as to readjust teaching programs or approaches and above all to give each student proper advice based on the individual-related feedback data.

Based on what has been discussed above the paper advocates that apart from having the test paper timely available, the authorities concerned provide detailed score reporting (Bailey, 1996), give feedback to teachers and students so that meaningful change can be effected (Heynelman & Ransom, 1990) and make sure teachers and administrators are involved in different phases of the testing process because they are the people who will have to make changes (Shohamy, 1992).

4. Investigation and reflection on the washback of testing to ELT

4.1 The influence of NMET and CET-4/CET-6 on ELT

Based on the questionnaires, classroom observation and interviews with senior high school students, college students, teachers and administrators concerned, a comparison is made about the washback effects of NMET and CET-4/CET-6 on ELT in Table 2:

<table>
<thead>
<tr>
<th>Test</th>
<th>NMET</th>
<th>CET-4</th>
<th>CET-6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Students’ affects to the test</td>
<td>Over 90% of the candidates are highly-motivated in the test preparation. The pre-test anxiety is high with hardly ever chance to have another try.</td>
<td>Not every sophomore is actively prepared for CET-4. With the chance to have another try, students’ pre-test anxiety is not so high as that to NMET. But some under-average students lose confidence in study after successive failures in trying to</td>
<td>Not all CET-4 passers are to sit for CET-6. The pre-test anxiety is nearly identical to that of CET-4, failure in the test has less impact on their</td>
</tr>
</tbody>
</table>
Teachers’ affects to the test

- Teachers are actively engaged in preparing students for NMET, and are all heavily stressed.
- Teachers’ affects vary with individual outlook of teaching, experiences and varied assessment institution, yet are mostly stressed.
- Essentially identical to that of CET4.

Positive effect on teaching and learning

- It motivates students to English learning and the course draws students’ attention, which, in turn, facilitates teaching.
- It is set as a phase goal by most college students and hence constitutes their instrumental motive. The course gains the students’ attention, which, in turn, facilitates teaching.
- It serves as another phase goal and hence the motivation for academically good students.

Negative effect on teaching and learning

- Competitive study strains the students and causes anxiety, uneasiness & fear on the part of students and test-oriented teaching on the part of teachers.
- Engaged in sample tests or cramming up CET-4 word list, some students give up regular English study. Some CET-4 passers stop learning English and some colleges designate weeks for pre-test tutoring, affecting course content and causing loss of instructional time, etc.
- To pass CET-6, some CET-4 passers give up regular English study and neglect fostering skills that require complex thinking and problem-solving ability.

Effect on the content, method, degree, depth of teaching

- Testing directly affects senior high school English teaching and learning in the content, method, degree & depth of teaching.
- Testing affects college English teaching and learning in the content, method, degree and depth of teaching.
- Testing affects college English teaching and learning but not as strong as that of CET-4.

Related research on testing

- The two national core journals open up the testing column for teachers nationwide to analyze every NMET and probe the enlightenment or washback effect on teaching and learning.
- Except occasional theoretical theses on testing, the14 core foreign language journals of the nation hardly leave any regular space for teachers to probe the washback effect on college ELT, or for empirical research discussion though there is said to be some break.
- The same as that of CET-4.

4.2 Testing as a guidance to ELT

As is proved in practice, the notion of teaching reform finds a quicker and easier way to individual school, public or private, via testing than any other means. Since NMET and CET-4/6 are the main concerns in general English education, test designers are able to transmit any message of teaching reform by adjusting or modifying the test content or format, while teachers, by analyzing the test, can catch the message and then make corresponding readjustment. For example, in CET-4/CET-6 of 2003, the dramatic reduction in the multiple choice of structure-related items in VS, and the increase in testing on comprehension of involved sentences in RC, SA, etc. signalize that grammar teaching should be updated from structure recognition to structure application. Likewise, the innovative form of writing also embodies the notion of skill-fostering inclination. Teachers, by analyzing the test paper if available, would grasp the message and readjust their foci of teaching by diverting their attention from giving students endless multiple choice of structure-recognition exercises to designing authentic reading and writing tasks to practice. In short, they would follow the guidance conveyed by the test to improve students’ performance and develop their language competence.
4.3. Research on testing as a task for both test-designers and teachers

The comparison in Table 2 confirms the hypotheses that key tests yield a multidimensional influence on ELT. Hence, research into such tests is not only a task for researchers and test-composers, but also a must for average English teachers. Table 2, item 6 particularly compares the differences in the issue of testing research to denote that NMET enjoys a regular column of testing research set aside by the two national core journals, while the 50,000 college English teachers (of the general English course) of the nation hardly any access to the 14 core journals of foreign language to air their view about testing or teaching practice. Bailey (1996) holds that if teachers understand the rationale, the purpose and directions of test paper-composition, it will greatly help their teaching practice, encourage them to get involved in teaching reform and promote them in their career-development. As reflective practitioners, the college English teachers, apart from their teaching job, need to get engaged in related research work, of which, testing is an indispensable part, the probe into which needs empirical evidence from teachers. Just as Alderson and Wall (1993) pointed out, “While claims of washback and its effects, both positive and negative, are numerous in educational literature, little empirical evidence has been provided to support the argument that tests do influence teaching; that is, that washback actually exists.” Therefore, the paper advocates the national core journals to open up some space to engage teachers in the empirical research to amend the lack and enhance their competence in both teaching and research.

Conclusion: The investigation into NMET and CET-4/CET-6 demonstrates that in China important tests do affect both what and how teachers teach, though the effect differs in degree and kind from teacher to teacher. To promote positive washback, the paper appeals to a more detailed score reporting on levels of students’ performance and areas of difficulty to be provided for teachers to get enlightenment of the guiding message, respond to current reform calls and follow closely to the syllabus so as to avoid blindness in their job, neither getting trapped in test-oriented teaching nor being indifferent to students’ concerns. Finally, the paper appeals to the 14 core journals to assist college English teachers in their career-development by opening up some forum to engage them in related research and share their view.

References
at Manoa,


[8] Huang Da-yong Yang Bing-jun, 2002 A Survey on the washback of language testing [J] Foreign Language Teaching and Research No.4
