

Washback of Testing on Teaching

Karmen Štumberger

Two terms are generally used in connection with the effect of tests on teachers, teaching content and society in general: wash-back and impact.

Wash-back

The term wash-back (also known as backwash) refers to the influence of language testing on teaching and learning. As McEwen has put it back in 1995, “what is assessed becomes what is valued, which becomes what is taught “. The assumption is that teachers are influenced by the knowledge that their students are planning to take a certain test. Therefore, they make changes in their teaching methodology and lesson content to help their students to perform better on the test (Taylor, 2005).

Wash-back is perceived as being either negative or positive. Testing can exert either a positive or negative influence on teaching. Negative wash-back occurs when the test’s content or format is based on a narrow definition of language ability and as such restrains teaching and learning process (Taylor, 2005). A rather extreme example of a negative wash-back was given by Davies et al. (1999): “If the skill of writing is tested only by multiple choice items then there is great pressure to practice such items rather than to practice the skill of writing itself “. The pressure that an upcoming test puts on teachers and learners can also be regarded as a negative wash-back. Positive washback, however, occurs when the testing procedure encourages good teaching practice, for example, the introduction of an oral proficiency test may promote the teaching of speaking skills, or practice of some examination tasks may improve the result on the test. In addition to that, creative and innovative testing can result in a syllabus change and the development of new teaching materials.

Impact

In the past two decades we have witnessed that testing can have consequences beyond just the classroom (Taylor, 2005). Test results may have a significant impact on the career or life chances of individual test takers (e.g. employment opportunities) or on society more widely. They may also influence educational systems (e.g. curriculum planning), e.g. they may lead to production and publishing of test preparation materials, course books and organisation of test preparation courses, for example, aviation community benefited from the publication of Aviation English Course books that proved to be good resource books for teaching and learning English needed by pilots and controllers.

Assertions about the effects of wash-back and impact have often been based on assumptions rather than on measurable evidence. So, researchers have been called upon to conduct more empirical studies on the impact of a particular test. They should provide

evidence that students have learned more or better because they have studied for a particular test.

Aviation English Test

Traditionally, tests have come at the end of the teaching and learning process. With the widespread expansion of high-stakes examinations, the direction seems to have been reversed. Testing can now come first in the whole process (Cheng & Curtis, 2004).” This is also true of the teaching and testing in the aviation sphere. Language assessment in aviation was introduced as a response to the findings that several accidents were, at least partly, caused by miscommunication between pilots and controllers. The introduction of the requirement for assessment of language proficiency brought about the need for training and hence for the design of teaching materials. The purpose of language testing in aviation is to ensure the operational level of language proficiency in the language used for radiotelephony communications of pilots and controllers to achieve safe and reliable communications. The language used in international radiotelephony communications is English.

As our colleague, Philip Shawcross stated, “proficiency testing in an aviation context is truly high-stakes, because its outcome directly engages the real world in terms of safety and career”. Aviation English tests are high-stakes tests which may influence lives and careers of pilots and air traffic controllers. The required minimum language proficiency in aviation is the operational level, L4. But, as it is evident from the open job vacancies, some companies require the extended or even the expert level. This is debatable and I am sure that the PRICE Study Group didn’t have this in mind when developing the language proficiency requirements. Yet, we should not turn a blind eye to the fact that language proficiency endorsement is used as a selection criterion.

Test driven instruction

The fact is that teachers want students to perform well on the tests and teach knowledge and skills assessed by the test. They make changes in their teaching to help their students to perform better on a test. Despite some “reservations” regarding measurement-driven instruction, it has been acknowledged that with high-stakes testing, tests should drive teaching. As suggested by Cheng and Curtis in 2001, an overlap between the content and format of the test and the content and format of the curriculum should be encouraged.

Positive wash-back is often associated with test validity; that is, the degree to which the test is measuring what it claims to measure. According to ICAO language requirements, aviation English tests should measure communicative language skills in work-related contexts. If the teaching pursues the same aim, it should result in the improvement of language proficiency.

Let us examine two examples of a positive wash-back on teaching and learning.

1. Refresher courses for air traffic controllers

The first round of English language testing of Slovene air traffic controllers was completed in 2009. The results revealed that 24 % of all controllers (26 controllers out of 111) failed the test, 58 % (65 controllers) managed to reach the operational level and only 18 % (20 controllers) could boast of the extended level. The majority of the candidates who did not succeed in the test failed the listening part of the test. The reason for bad results was not the lack of proficiency of our controllers but the lack of preparation. The examiners believed that familiarisation with the test format and completion of the sample test will suffice. We have to bear in mind that these controllers had not undergone EL training since their initial training. The controllers, however, did not bother to take the sample test, or checked the web page the night before the test date when the server was accidentally down or had a look at it an hour before the test and only managed to go through the first half of the tasks. Everybody was appalled by the results. In the end, after one or more re-takes, everybody succeeded in achieving the operational level. When the third round of testing for L4 candidates and the second round for L 6 started in 2014, the candidates were not younger and no EL training had taken place since the first trial despite my recommendations. Yet, we were confident that everybody is familiar with the test and that there will be no problems. The first few testing sessions, however, showed the opposite. 6 test takers again failed the test. The training organisation of the national air traffic services provider reacted at once, contacted a language school which organised refresher courses. The school was instructed to examine the Manual on the Implementation of ICAO Language Proficiency Requirements (Doc 9835) and the test. Besides dealing with aviation topics, standard grammar mistakes, the teachers introduced tasks used in the test (e.g. picture description, etc.). The most original one was the listening task in which the candidates had to listen to radiotelephone conversations and answer associated questions which presented the hardest nut to crack for test takers. To make the task as realistic as possible, the teacher brought the computers and the participants had to write the answers into spaces provided practicing simultaneously their typing skills.

The outcomes of the course were better than expected. From introduction of the refresher courses on only one candidate has failed the exam. Moreover, the controllers were satisfied with the course, they found it useful, learned for the test, tried to achieve a higher grade and their overall performance on both the listening test and the speaking test improved. Since the test design was based on the analysis of language needs of air traffic controllers, and the test assesses the test-takers' ability to communicate effectively in work-related conversations, we may presume that their language performance in aeronautical communication has improved. This leads to the achievement of a higher goal, safer and more reliable communication. In this example, the test influenced teaching and learning and had a positive impact on test results.

2. Achievement of a higher grade

In the assessment of English language proficiency of pilots, the failure rate was much lower. In the last couple of years, however, several pilots expressed their dissatisfaction with the operational level of English Language Proficiency and the wish to sit for the test before the expiry of the language endorsement. The reason was recruitment adds requiring a higher EL proficiency grade. These candidates received a L4 in maximum tree categories. They were told what is expected from a L5 candidate in a category where they demonstrated a lower level. Further, they were encouraged to read aviation related articles, watch aviation series, listen to aviation recordings, revise basis grammatical structures and examine the rating scale. After a couple of weeks of self-study they took the test again and demonstrated improved proficiency. A similar thing happened with two candidates wishing to demonstrate language proficiency at L6. Here too, familiarity with the expected performance at the desired level was crucial.

In cases where candidates are motivated to improve their language ability because of job opportunities, familiarity with the rating scale and the performance at that level is of great importance. In these cases too, testing exerted a positive influence on learning.

To conclude, teachers should be aware of the effects of testing and try to foster beneficial wash-back. Their understanding of the test and the rationale behind its design shall be reflected in the content and method of teaching. In addition, teachers should establish and maintain cooperation with the testing agencies. Teacher, testers and regulator should be aware of the connection between testing and training to ensure the improvement in language proficiency.

References

Cheng, L. and Y. Watanabe.2004. *Washback in Language Testing*. New York: Routledge.

Shawcross, P. "What do we mean by the 'washback effect' of testing? ". Retrieved from: <http://www.icao.int>

Taylor, L. 2005. "Washback and Impact". *ELT Journal* 59 (2): 154-155.