

A review of washback and its pedagogical implications

Yi-Ching Pan*

The university of Melbourne, Australia

Received 4 February 2009

Abstract. The way in which examinations influence teaching and learning is commonly described as “washback” or “backwash”. A number of definitions have been proposed for the term “washback” throughout the published research and literature on language testing. This study starts with a focus on the various definitions of backwash or washback. Next, it examines the similar concept terms defined by other researchers. By reviewing the variety of definitions, the researcher’s own view of washback will be reached. Comes after that are the explorations of different types of washback. The studies ends with drawing pedagogical implications for EFL teachers.

1. Introduction

It is a common belief that testing affects teaching and learning, as stated by Alderson and Wall (1993 [1]) that “tests are held to be powerful determiners of what happens in classroom”. The way in which examinations influence teaching and learning is commonly described as “washback” or “backwash”. A number of definitions have been proposed for the term “washback” throughout the published research and literature on language testing. This study starts with a focus on the various definitions of backwash or washback. Next, it examines the similar concept terms defined by other researchers. By reviewing the variety of definitions, the researcher’s own view of washback will be reached. Comes after that are the explorations of different types of washback. The studies ends with drawing pedagogical implications for EFL teachers.

2. Definitions of Washback

The notion of “washback” is prevalent in language teaching and testing literature, but it is seldom found in dictionaries. Some writers used the term “washback” while others preferred “backwash” to describe the effects or influences brought by tests or examinations. Below, the definitions by various researchers are arranged under the groupings of (a) backwash or (b) washback.

2.1. Backwash

- Hughes (1989:1): “The effect of testing on teaching and learning” is known as backwash.

- Spolsky (1994 [2]): The concept of backwash deals with the unforeseen side-effects of testing and not to the intended effects when the primary goal of the examination is the control of curricula.

- Biggs (1995 [3]): Backwash refers to the fact that testing controls not only the curriculum

* E-mail: y.pan@pgrad.unimelb.edu.au

but also teaching methods and students' learning strategies.

2.2. Washback

- Alderson & Wall (1993 [1]): Washback compels "teachers and learners to do things they would not necessarily otherwise do because of the test".

- Messick (1996 [4]): Washback is described as "the extent to which the introduction and the use of a test influences language and teachers to do things they would not otherwise do that promote or inhibit language learning".

- Bailey (1996 [5]): Washback is the "influence of testing on teaching and learning."

- Shohamy, et al. (1996 [6]): Washback is delineated as "the connections between testing and learning".

- Pearson (1998 [7]): "Public examinations influence the attitudes, behaviours, and motivation of teachers, learners, and parents, and because examinations often come at the end of a course, this influence is seen working in a backward direction, hence the term, washback".

- Cheng (2005 [8]): Washback indicates "an intended or unintended (accidental) direction and function of curriculum change on aspects of teaching and learning by means of a change of public examinations".

3. Similar concept terms to washback

In addition to "backwash" and "washback", researchers used other similar terms stated below to investigate the phenomena of the influences or effects of tests on the educational field.

a) Test impact (Andrews, 2004 [9]; McNamara, 2000 [10]; Wall, 1997 [11]; Bachman & Palmer, 1996 [12])

Some researchers have argued that tests can have more far-reaching effects in the educational world than just in the language

classroom. Bachman & Palmer (1996 [12]) used the term "test impact" to refer to the effects that tests have on individuals (teachers and students) or educational systems and on the society at large. Wall (Wall, 1997 [11]) held a similar view by stating that "Test impact refers to any of the effects that a test may have on individuals, policies or practices within the classroom, the school, the educational system, and society as a whole". McNamara (2004 [10]) claimed that "Tests can also have effects beyond the classroom. The wider effect of tests on the community as a whole, including the school, is referred to as test impact". Andrews (2004 [9]) used "test impact" to describe "the effects of tests on teaching and learning, the educational system, and the various stake holders in the education process".

b) Systemic validity (Fredericksen & Collins, 1989 [13])

Systemic validity refers to the effects of instructional changes brought about by the introduction the test into an educational system as stated tests induce "in the education system curricular and instructional changes that foster cognitive skills that the test is designed to measure".

c) Consequential validity (Messick, 1989, 1996 [4,14])

Consequential validity encompasses concepts ranging from the uses of tests, the impacts of testing on test takers and teachers, the examination of results by decision makers, and the potential misuse, abuse, and unintended usage of tests. In other words, consequential validity implies that tests have various influences both within and beyond the classroom. In other words, consequential validity refers to the societal implications of testing that are only one facet of a broader, unified concept of test validity.

Measurement-driver instruction (Shohamy 1992 [15])

Shohamy contended that "the use of external tests as a device for creating impact on

the educational process is often referred to as the washback effect or measurement-driven instruction”.

d) Curriculum-alignment (Shohamy et al 1996 [6])

Shohamy et al. defined curriculum-alignment as “the curriculum is modified according to test results”.

e) Washback validity: (Morrow 1986 [16])

“In essence, an examination of washback validity would take testing researchers into the classroom in order to observe the effect of their tests in action.” In other words, washback validity deals directly with the extent to which the test meets the needs of students, educators, researchers, administrators of tests, and anyone who uses the test results in the future. In addition, washback validity refers to the value of the relationship between the test and any associated teaching

4. Definition of washback in this study

After reviewing definitions of washback, the term can be defined according to two major perspectives: one at a narrower view within the classroom at a micro level, and the other at a wider and more holistic view beyond the classroom at a macro level.

As suggested by Bachman & Palmer (1996 [12]), washback, at a macro level, refers to the extent to which a test influences within the society, ranging from government policy-making, school administration, publishing, and general opportunities, to parents’ expectations of their children. At a micro level, washback refers to the extent to which a test influences within the classroom, mainly in the change or innovation of curricula and teachers’ methodologies and the influence of students’ learning. Bailey (1996 [5]) used the phrase “washback to the learners” to indicate the effects of test on students, and “washback to the programme” to indicate effects of test on teachers, administrators, curriculum developers, counselors, etc.

To summarize, the narrower definition of washback focuses on the effects that a test has on teaching and learning. The wider or more holistic view of washback (also defined as test impact) looks beyond the classroom to the educational systems and society at large. All in all, tests can have “significant impact not only on individuals but also on practices and policies—in the classroom, the school, the educational system and in society as a whole” (Wall 2005 [17]).

In this study, a broader interpretation of washback will be adopted: washback at a macro level beyond the classroom to investigate the test washback in the school, the educational system and society as a whole, and washback at a micro level within the classroom to investigate the test washback in the classroom, that is, the washback effect of teaching and learning.

Types of Washback

Generally, washback can be analysed according to two major types: positive and negative, depending on whether it has a beneficial or harmful impact on educational practices (Hughes, 1989). This section explores positive and negative washback in terms of both the classroom setting and the educational/political system.

5. Positive Washback

5.1. Classroom setting

Teachers and learners will be motivated to fulfil their teaching and learning goals (Anderson & Wall, 1993 [1]).

Good tests can be utilized and designed as beneficial teaching-learning activities so as to encourage a positive teaching-learning process (Pearson, 1988:107).

A creative and innovative test can quite advantageously result in a syllabus alteration or a new syllabus (Davis, 1985 [18]).

5.2. Educational/societal system

Decisional makers use the authority power of high-stakes testing to achieve the goals of teaching and learning, such as the introduction of new textbooks and new curricula (Shohamy, 1992 [15]; Wall & Alderson 1993 [1]; Cheng; 2005 [8]).

Tests are encouraged to promote the idea of lifelong learning and encourage people to learn English (Language Testing and Training Centre, 2008).

5.3. Negative Washback

Classroom setting

The test will lead to the narrowing of content in the curriculum. What students have learned is test language, instead of total phases of understanding (Shohamy, 1992 [15]).

Teachers tend to ignore subjects and activities that are not directly related to passing the exam, and tests accordingly alter the curriculum in a negative way (Vernon, 1956 [18]).

The tests may well fail to create a correspondence between the learning principles and/or the course objectives to which they should be related (Cheng, 2005 [8]).

Many teachers detailed high anxiety, fear and pressure to cover the material, as they felt that their job performance was assessed by students' test scores (Shohamy, 1996 [6]).

Educators experienced negative reactions to the stress brought about by public displays of classroom scores. Inexperienced teachers felt a

greater degree of anxiety and pressure for accountability than did teachers with more experience (Fish, 1988 [19]).

“Testing programs substantially reduce the time available for instruction, narrow curricular offerings and modes of instruction, and potentially reduce the capacities of teachers to teach content and to use methods and materials that are incompatible with standardized testing formats” (Smith, 1991 [20]).

An increasing number of paid coaching classes are set up to prepare students for exams, but what students learn are test-taking skills rather than language learning activities (Wiseman, 1961 [21]).

Measurement-driven instruction will definitely result in cramming, narrowing the curriculum, focus of attention on those skills that are most relevant to testing, placement of constraints on teachers' and students' creativity and spontaneity, and disparage the professional judgment of educators (Madaus, 1988 [22]).

5.4. Educational/societal system

Decision makers overwhelmingly use tests to promote their political agendas and to seize influence and control of educational systems (Shohamy, 1996 [6]). Tests are used as a “lever” for change.

To present a clear view of positive and negative washback at both micro-level (classroom settings) and at macro-level (educational and societal system), Tables 1 and 2 is presented below for classification.

Table 1. Summary of Positive washback

Positive Washback	
Classroom settings	<ol style="list-style-type: none"> 1. Tests induce teachers to cover their subjects more thoroughly, making them complete their syllabi within the prescribed time limits. 2. Tests motivate students to work harder to have a sense of accomplishment and thus enhance learning. 3. Good tests can be utilized and designed as beneficial teaching-learning activities so as to encourage positive teaching-learning processes.
Educational/societal system	Decision makers use the authority power of high-stakes testing to achieve the goals of teaching and learning, such as the introduction of new textbooks and new curricula.

Table 2. Summary of Negative washback

Negative Washback	
Classroom settings	<ol style="list-style-type: none"> 1. Tests encourage teachers to narrow the curriculum and lose instructional time, leading to “teaching to the test.” 2. Tests bring anxiety both to teachers and students and distort their performance. 3. Students may not be able to learn real-life knowledge, but instead learn discrete points of knowledge that are tested. 4. Cramming will lead students to have a negative positive toward tests and accordingly alter their learning motivation.
Educational/societal system	Decision makers overwhelmingly use tests to promote their political agendas and to seize influence and control of educational systems.

To summarize, in terms of the classroom setting at a micro level, the positive washback integrates meaningful and innovative learning activities in teachers' educational methodologies, and thus educators will devote more attention to students' intentions, interests, and choices. Students at the same time will be encouraged and motivated to work harder. On the other hand, the negative washback is that teachers will usually teach to the test, narrow the curriculum and only focus on what will be tested. Moreover, cramming will be the washback brought by measurement-driven tests, even though there is an ongoing debate as to whether cramming is positive or negative washback. In terms of educational setting, the positive washback is that the authority can use the test to attain its goal of teaching and learning. However, the negative washback is

that the authority uses that goal to control and obtain the power of the academic system that will usually place undue pressure and anxiety on school staffs, teachers and even students. In other words, the washback on the side of the educational setting is one coin with two sides, depending on the stakeholder's point of view.

6. Pedagogical Implications

By analyzing the possible positive and negative washback that tests might bring about at micro and macro levels, it seems that teachers play an important role in fostering different types of washback. In other words, the beliefs of the teachers are a critical factor in determining the washback effect. For example, a test, on one hand, will encourage some

teachers to think it's important to plan their curricula carefully to meet the needs of the test, but on the other hand, other teachers may think that tests force them to teach what they don't find suitable or appropriate for students. Spratt (2005) has stated that the teacher plays a significant role in determining the types and intensity of washback, and thus, teachers have become the sources of promoting positive washback. Chapman and Snyder J. (2000:462) have expressed a similar view by stating that "its is not the examination itself that influences teachers' behavior, but teachers' beliefs about those changes". As Watanbee (2005) suggested, teachers should be provided with in-service training and be familiar with a wide range of teaching methods.

Tests sometimes are used by schools or school administrations as a "lever" to introduce the innovation of new curricula, but it may change the format of what teachers instruct, not foster an in-depth change of teaching methodologies as a whole. As Wall (2005:283) stated, "examinations cannot influence teachers to change their practices if they are not committed to the new ideas and if they do not have the skills that will enable them to experiment with, evaluate and make appropriate adjustments to new methods". In other words, teachers themselves must conduct the changes and teachers need to have the necessary skills to adapt the changes. Again, teachers play a very crucial role in promoting positive washback or hindering negative washback.

To conclude, there are two major perspectives that teachers should bear in mind. If we are the ones who make the tests, we should try to make a match between what is tested and what is taught by using more direct testing, making sure the test is known by students. Tests are one factor that will lead the teacher to "teach to the test", and what students learn might be discrete points of language, not the communicative part of language they need in real life. To remedy this, it is desirable to use authentic and direct tests (Bailey, 1996). If we

are responsible for helping students pass the test, we should try our best to learn more teaching methodologies by taking more training courses, engaging in peer observations and utilizing the tests to enhance students' learning while at the same time not inhibiting students' motivation by cramming too much. As teachers, "we may have limited power to influence high-stakes national and international examinations, but we do have tremendous power to lead students to learn, to teach them language and how to work with tests and test results." (Bailey, 2005) All in all, it's the teacher who has the most power to turn it into positive or negative washback.

References

- [1] C. Alderson, D. Wall, Does washback exist? *Applied Linguistics*, 14 (1993) 115.
- [2] B. Spolsky, *The examination of classroom backwash cycle: Some historical cases*, in Nunan, D, Berry, V and Berry, R. (Eds) Bringing about change in language education, University of Hong Kong, Dept. of Curriculum Studies, Hong Kong, 1994.
- [3] J.B. Biggs, Assumptions underlying new approaches to assessment, *Curriculum Forum*, 4 (1995) 1.
- [4] S. Messick, Validity and washback in language testing. *Language Testing* 13 (1996) 241.
- [5] K.M. Bailey, Working for washback: A review of the washback concept in language testing, *Language Testing*, 13 (1996) 257.
- [6] E. Shohamy, S. Donitsa-Schmidt, L. Ferman, Test impact revisited: Washback effect over time, *Language Testing* 13 (1996b) 298.
- [7] L. Pearson, Tests as levers of change (or "putting first things first"). In D. Chamberlain & R. Baumgartner (Eds.), *ESP in the classroom: Practice and evaluation ELT Documents #128*, (pp. 98-107), Modern English Publication in association with the British Council, London, 1988.
- [8] L. Cheng, *Changing Language Teaching Through Language Testing: A Washback Study*, Cambridge University Press, Cambridge, 2005.
- [9] S. Andrews, Washback and Innovation, In L. Cheng, Y. Watanabe & A. Curtis (Eds.), *Washback in Language testing* (pp. 37-50), Mahwah, New Jersey: Lawrence Erlbaum Associates, Publishers, 2004.
- [10] T. McNamara, *Language Testing*, Oxford University Press, Oxford, 2000.

- [11] D. Wall, Impact and washback in language testing, In C. C. & D. Corson (Eds.), *Encyclopaedia of language and education, Language Testing and Assessment 7* (1997) 291.
- [12] F. Bachman, S. Palmer, *Language testing in Practice*, Oxford University Press, Oxford, 1996.
- [13] R. Fredericksen, A. Collins, A system approach to educational testing *Educational Researcher* 18 (1989) 27.
- [14] S. Messick, Validity. In R. Linn (Ed.), *Educational Measurement* (pp. 13-103), Macmillan, New York, 1989.
- [15] E. Shohamy, Beyond proficiency testing: A diagnostic feedback testing model for assessing foreign language learning, *The Modern Language Journal*, 76 (1992) 513.
- [16] K. Morrow, The evaluation of tests of communicative performance, In E. Portal (Ed.), *Innovations in Language Testing: Proceedings of the IUS/NFER Conference* (pp. 1-13), NFER/Nelson, London, 1986.
- [17] D. Wall, *The impact of high-stakes examinations on classroom teaching*, Cambridge University Press, Cambridge, 2005.
- [18] P.E. Vernon, *The measurement of abilities*, University of London Press, London, 1996.
- [19] J. Fish, *Responses to mandated standardised testing. Unpublished doctoral dissertation*, University of California, Los Angeles, 1988.
- [20] M.L. Smith, Put to the test: The effects of external testing on teachers, *Educational Researcher* 20 (1991) 8.
- [21] S. Wiseman, *Examinations and English education*, Manchester University Press, England, 1961.
- [22] G.F. Madaus, The influence of testing on the curriculum. In L. N. Tanner (Ed.), *Critical issues in curriculum: Eighty-seventh yearbook of the National Society for the study of education* (pp. 83-121), University of Chicago Press, Chicago, 1988.

Điểm tư liệu về khái niệm "kênh hồi đáp" và những gợi ý sư phạm

Phan Di Tịnh

Đại học Melbourne, Australia

Phương thức mà các kỳ thi ảnh hưởng tới việc dạy và học thường được miêu tả như là “kênh phản hồi”. Đã có nhiều định nghĩa về “kênh phản hồi” được đề xuất trong các nghiên cứu và tư liệu bàn về kiểm tra - đánh giá ngôn ngữ. Bài viết này bắt đầu bằng việc nhấn mạnh vào nhiều định nghĩa khác nhau về “kênh phản hồi”. Tiếp theo, chúng tôi khảo sát một số thuật ngữ có cùng khái niệm tương tự trong một số nghiên cứu khác. Với việc điểm lại một số định nghĩa khác nhau, tác giả bài viết đưa ra quan điểm riêng của mình về vấn đề này. Tiếp theo là những khám phá về các loại hình “kênh phản hồi”. Kết thúc bài viết là những đề xuất mang tính sư phạm dành cho các giáo viên dạy tiếng Anh như một ngoại ngữ.