

Literature Review on Washback Effect of the Language testing

Zhang Chaofeng

Lijiang College of Guangxi Normal University, Guilin 541006, China

Keywords: Washback Effect, Language Testing, Teaching and Learning

Abstract: as a term commonly used in applied linguistics , washback effect is defined as the influence of the language testing on teaching and learning .This paper focus the mind on conducting a literature review toward washback effect through analyzing its definition, Classification, mechanisms and relevant study in China and other countries. The aim of this paper is 1) to expatiate the various definitions of Washback, find a general definition of the term "Washback", 2) to logically classify the effects of washback according to the result of the researchers, 3) to illustrate the mechanisms of the washback effect through the finding of the researchers, 4) to investigate the research status on wachback effect in China and other countries according to the result of the researchers.

1. Definition of Washback Effect

According to Liying Cheng & Andy Curtis, although washback is a term commonly used in applied linguistics today, it is rarely found in dictionaries. However, the word backwash can be found in certain dictionaries and is defined as “the unwelcome repercussions of some social action” by the *New Webster’s Comprehensive Dictionary*, and “unpleasant after-effects of an event or situation” by the *Collins Cobuild Dictionary*. The negative connotations of these two definitions are interesting, as they inadvertently touch on some of the negative responses and reactions to the relationships between teaching and testing, which we explore in more detail shortly (Cheng & Curtis, 2004).

Wash back (Alderson & wall, 1993) or backwash (Biggs, 1995, 1996) here refers to the influence of testing on teaching and learning. The concept is rooted in the notion that tests or examinations can and should drive teaching, and hence learning, and is also referred to as measurement-driven instruction (Popham, 1987). In order to achieve this goal, a “match” or an overlap between the content and format of the test or the examination and the content and the format of the curriculum (or “curriculum surrogate” such as the textbook) is encouraged. This is referred to as *curriculum alignment* by Shepard (1990, 1991b, 1992, 1993). Although the idea of alignment--matching the test and the curriculum—has been described by some as “unethical,” and threatening the validity of the test (Haladyna, Nolen, & Haas, 1991, p. 4; Widen, O’Shea, & Pye, 1997), such alignment is evident in a number of countries, for example, Hong Kong (Cheng, 1998a; Stecher, Barron, Chun, Krop, & Ross, 2000), such alignment, in which a new or revised examination is introduced into the education system with the aim of improving teaching and learning, is referred to as *systemic validity* by Frederiksen and Collins (1989), *consequential validity* by Messick (1989, 1992, 1994, 1996), and *test impact* by Batchman and Palmer (1996) and Baker (1991).

Wall (1997) distinguished between test impact and test washback in terms of the scope of the effects. According to Wall, 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 or society as a whole” (see Stecher, Chun, & Barron,1999), whereas washback (or backwash) is defined as “the effect of tests on teaching and learning” (Wall, 1997: 291).

Although different terms are preferred by different researchers, they all refer to different facets of the same phenomenon—the influence of testing on teaching and learning.

2. The Classification of Washback Effect

The potentially bidirectional nature of washback has been recognized by, for example, Messick (1996), who defined washback as the “extent to which a test influences language teachers and learners to do things they would not necessarily otherwise do that *promote or inhibit* [emphasis added] language learning” (cited from Alderson & Wall, 1993:117). Wall and Anderson also noted that “tests can be powerful determiners, both positively and negatively, [emphasis added] of what happens in classrooms” (Alderson & Wall, 1993:117; Wall & Anderson, 1993: 41).

According to Cheng & Curtis, whether the effect of testing of testing is deemed to be positive or negative should also depend on who it is that actually conducts the investigation within a particular education context, as well as *where*, the school or university contexts, *when*, the time and duration of using such assessment practice, *why*, the rationale, and *how*, the different approaches used by different participants within the context. (Liyang Cheng & Andy Curtis, 2004)

2.1 Negative Washback Effect

Test in general and perhaps language tests in particular, are often criticized for their negative influence on teaching-so-called “negative washback” which has long been identified as a potential problem. For example, nearly 50 years ago, Veron (1956) claimed that teachers tended to ignore subjects and activities that did not contribute directly to passing the exam, and that examinations “distort the curriculum”. Wiseman (1961) believed that paid coaching classes, which were intended for preparing students for exam, were not a good use of the time, because students were practicing exam techniques rather than language learning activities.

More recently, Alderson and wall (1993) referred to negative washback as the undesirable effects on teaching and learning of a particular test deemed to be “poor”. According to Liyang Cheng & Andy Curtis, Alderson and Wall’s “poor” here means “something that the teacher or learner does not wish to teach or learn.” The tests may well fail to reflect the learning principles or the course objectives to which they are supposedly related. In reality, teachers and learners may end up teaching and learning toward the test, regardless of whether or not they support the test or fully understand its rationale or aims. (Cheng & Curtis, 2004)

In general education, Fish (1998) found that teachers reacted negatively to pressure created by public displays of classroom scores, and also found that relatively inexperienced teachers felt greater anxiety and account ability pressure than experienced teachers, showing the influence of factors such as age and experience.

2.2 Positive Washback Effect

According to Cheng & Curtis (2004), like most areas of language testing, for each argument in favor or opposed to a particular position, there is a counterargument. There are, then, researchers who strongly believe that it is feasible and desirable to bring about beneficial changes in teaching by changing examinations, representing the “positive washback” scenario, which is closely related to “measurement-driven instruction” in general education. In this case, teachers and learners have a positive attitude toward the examination or test, and work willingly and collaboratively toward its objectives.

According to Pearson (1998), a test’s washback effect will be negative if it fails to reflect the learning principles and course objectives to which the test supposedly relates, and it will be positive if the effects are beneficial and “encourage the whole range of the quality of desired changes”. According to Liyang Cheng & Andy Curtis, it is possible that research into washback may benefit from turning its attention toward looking at the complex cause of such a phenomenon in teaching and leaning, rather than focusing on deciding whether or not the effects can be classified as positive or negative. (Liyang Cheng & Andy Curtis, 2004)

3. The Mechanisms of Washback

Bailey (1996) cited Hughes’ trichotomy (1993) to illustrate the complex mechanisms through

which washback occurs in actual teaching and learning environments (see Table 1). Hughes (1993) explained his model as follows:

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

Table 1 The Trichotomy Backwash Model

(A)	Participants—students, classroom teachers, administrators, materials developers and publishers, whose perceptions and attitudes toward their work may be affected by a test.
(B)	Process—any actions taken by the participants which may contribute to the process of learning
(C)	Products—what is learned (facts, skills, etc.) and the quality of the learning

Whereas Hughes focused on participants, process, and products in his model to illustrate the washback mechanism, Alderson and Wall (1993), in their Srilankan study, focussed on micro aspects of teaching and learning that might be influenced by examinations. Based on that study, they drew up 15 hypotheses regarding washback, which referred to areas of teaching and learning that are generally affected by washback. Andrews (1994, 1995) highlighted the complexity of the relationship between washback and curriculum innovation, and summarized three possible responses of educators in response to washback as follows:

- ✓ Fight it: the effort to replace examination with other sorts of selection processes and criteria.
- ✓ Ignore it: tests and exams have been seen as an occasional necessary evil, a dose of unpleasant medicine, the taste of which should be washed away as quickly as possible.
- ✓ Ues it: using washback to promote particular pedagogical goals.

4. Relevant Studies in China and Other Countries

4.1 Studies in China

Some researchers have studied washback effect from different points of view. Ping Yuan (2002) points out that attainment test can exert positive washback effect on foreign language teaching, prompt students' command of textbook knowledge and make for the improvement of students' integrating competence. According to Shaoshan Li (2005), the designers of language test, users of testing result as well as language teachers must pay high attention to the washback effect of test on teaching and learning. Luxia Qi (2004) indicates it is necessary to strength the after-effect study of high-risk examinations, thoroughly make clear the formation mechanism of after-effect and take positive measures specifically so as to reduce the negative influences of examinations. Dayong Huang & Bingjun Yang (2002) mainly introduces the finding of theoretical and empirical study on washback effect abroad as well as summarizes the research status in china. Hongxia Luo (2018) indicates the National English Contest produce the positive washback effect on the innovation of the teaching model as well as the evaluating system.

4.2 Studies in Other Countries

The task of the language test is conducting an objective, accurate and justice evaluation to the test-takers. After examination, the result of it will exert great influences on the teachers' teaching and students' learning. In the area of applied linguistics, such influences are called washback. How to analyze and evaluate the washback has become the important study topic for many linguists.

Pearson (1988), discussed the two facets of washbck, that is positive washbck and negative washback. His study has shown whether the washback effect is positive or not is dependent upon if the learning principles and course objectives that relate to test can be reflected. Hughes used the 3P model which is shown in table to illustrate the washback taking place in actual environment of the

teaching and learning. Alderson and Wall (1993) concluded that further research on washback is needed, and that such research must entail “increasing specification of the washback hypothesis”. One response to Alderson and Wall’s (1993) recommendation was a large-scale quantitative and qualitative empirical study, in which Cheng (1997, 1998a) developed the notion of “washback intensity” to refer to the degree of the washback effect in an area or a number of areas of teaching and learning affected by an examination.

5. Conclusion

The studies in this field have been published on several aspects by some Chinese researchers. They introduce the theory of washback effect, discuss its concept and characteristics, and predict its applied prospect of English teaching in China. Some articles even induct the model of washback effect. However, the author found that most of the studies are focused on the theoretical research, with few on the empirical research. Besides, although there are many studies related to the English learners, the specific studies according to two groups of learners with regard to test-takers and non-test-takers in some tests are rarely found.

According to the research on the washback effect in other countries, the research and analytical methods of study concerning or related to foreign language needs analysis are diverse and flexible. With the mainstream of empirical research methodology, the reliabilities and validities of these researches are still in question and urgently need to be developed.

References

- [1] Alderson, J. C., & Wall, D. Does washback exist? [J] *Applied Linguistics*, 1993, 14: 115-129.
- [2] Andrews, S. The washback effect of examinations: Its impact upon curriculum innovation in English language teaching [J]. *Curriculum Forum*, 1994, 4(1): 44-58.
- [3] Andrews, S. Washback or washout? The relationship between examination reform and curriculum innovation [A]. In D. Nunan, V. Berry, & R. Berry (Eds.), *Bringing about Change in Language Education* [C]. Hong Kong: University of Hong Kong, 1995: 67-81.
- [4] Bachman, L. F. & Pamler, A. S. *Language testing in practice*. [M] Oxford, England: Oxford University Press, 1996.
- [5] Bachman, L. *Fundamental consideration in language testing*. [M] Oxford, England: Oxford University Press, 1990: 94.
- [6] Baily, K. M. Working for washback: A review of the washback concept in language testing [J]. *Language Testing*. 1996, 13: 257-279.
- [7] Baker, E. L. Issues in policy, assessment, and equity [A]. Paper presented at the national research symposium on limited English proficient students’ issues: *Focus on evaluation and measurement* [C]. Washington, DC, 1991, September.
- [8] Biggs, J. B. Assumptions underlying new approaches to educational assessment [J]. *Curriculum Forum*, 1995, 4(2): 1-22.
- [9] Canale, M. & Swain, M. Theoretical bases of communicative approaches to second language teaching and testing [J]. *Applied linguistics*, 1980.
- [10] Cheng, L. The washback effect of public examination change on classroom teaching: *an impact study of the 1996 Hong Kong certificate of education in English on the classroom teaching of English in Hong Kong secondary schools* [P]. Unpublished doctoral dissertation, university of Hong Kong, Hong Kong, 1998.
- [11] Cheng, L. How does washback influence teaching? Implications for Hong Kong [J]. *Language and education*, 1997, 11: 38-54.

- [12] Dayong Huang & Bingjun Yang. The overview on language testing's washback effect [J]. *Foreign Language Teaching and Research*, 2002, (04).
- [13] Fish, J. *Responses to mandated standardized testing* [P]. Unpublished doctoral dissertation, university of California, Los Angeles, 1988.
- [14] Haladyna, T. M., Nolen, S. B., & Haas, N. S. Raising standardized achievement test scores and the origins of test score pollution [J]. *Educational Research*, 1991, 20(5): 2-7.
- [15] Hongxia Luo. The Positive washback effect on College English Teaching [J]. *Journal of Lanzhou University of Arts and Science (Social Science)*, 2018, (05).
- [16] Liying Cheng & Andy Curits. Washback or backwash: a review of the impact of testing on teaching and learning [A]. In Liying Cheng, Yoshinori Watanabe & Andy Curits, *Washback in Language Testing: Research Contexts and Methods* [C]. London, Lawrence Erlbaum. 2004.