Washback: Exploring what constitutes "good" teaching practices

Melissa Barnes

Monash University, Faculty of Education, McMahons Road, Frankston, VIC 3199, Australia

ARTICLE INFO

ABSTRACT

Realising the social consequences involved in language testing, many researchers have investigated "washback", or the influence of testing on teaching and learning. While it is widely acknowledged that the nature of washback is dependent on context (Cheng, 2005; Tsagari, 2011), the very definition of washback is problematic due to its reliance on what constitutes "good" teaching and learning practices, which can differ from one educational context to another. This article explores the attitudes, beliefs and teaching practices of four teachers teaching both TOEFL iBT preparation courses and general English classes in Vietnam with the aim to better understand how washback (positive or negative) is realised in these contexts. In this qualitative case study, the participating teachers were observed and interviewed in both their TOEFL iBT preparation and general English class. The study revealed that teachers' beliefs about what constitutes "good" teaching practices differed depending on the teaching context, or more specifically, the role and purpose of the course (test preparation versus general English).

© 2017 Elsevier Ltd. All rights reserved.

Washback, or "backwash," is used in applied linguistics to refer to the influence of testing on teaching and learning and has been described by researchers as a complex phenomenon consisting of numerous mediating factors. While most researchers agree that washback exists, they also acknowledge that there are varying degrees (Alderson & Hamp-Lyons, 1996; Spratt, 2005; Watanabe, 2004), intensity (Cheng, 2005; Green, 2007) and direction (Green, 2007) of washback. Washback is often characterised as either positive or negative and is recognized as playing an important role in the relationship between testing, teaching and learning.

Contrasting negative and positive washback, Taylor (2005) argues that negative effects occur "when a test's content or format is based on a narrow definition of language ability" while positive effects occur when the testing procedures encourage "good" teaching practices (p. 154). While Taylor's definition of positive and negative washback appears reasonable, what constitutes "good" teaching practices is often contextually defined. The reality is that teaching approaches evolve and change and therefore caution must be taken in defining phenomenon such as washback in connection with a preferred teaching paradigm (e.g. Communicative Language Teaching).

Similar to Bailey (1999), Taylor (2005) suggests that tests, which are not aligned with or "run contrary" to the principles and practices of Communicative Language Teaching (CLT), generate negative washback (p. 259). She goes on to argue that "it is unlikely that a test based on outmoded theoretical constructs will lead to positive washback. Since, in many parts of the world, a narrow view of linguistic competence has been replaced by a broader perspective on communicative competence" (p. 276).

E-mail address: melissa.barnes@monash.edu.

The views of Taylor (2005) and Bailey (1999) suggest that the way in which a test is designed, particularly if it aligns itself with either CLT or traditional teaching practices can determine whether the test's washback is positive or negative. This can then be realised through "good" teaching practices. A look into several prominent washback studies will attempt to uncover how the language testing community has come to define washback by the presence or absence of CLT.

1. CLT & high-stakes language testing

CLT emerged from a paradigm shift in the 1970's when linguists and language educators began to view language as a system for the expression of meaning rather than a system of syntactic rules. This shift began in the late 1960's when Hymes (1967; 1972) introduced a theory or model of communicative competence within a first language framework. Hymes distinguished between actual language used, language knowledge and the "ability for use." Using Hymes' first language model to inform their own, Canale and Swain (1980) adapted the model to suit a second language framework and created more specific domains of language knowledge: grammatical, sociolinguistic, and strategic competence. They argued that communicative competence consists of the interaction between knowing the rules of grammar (grammatical competence), the rules of language use (sociolinguistic competence) and the skills to overcome problems in communication (strategic competence). It was Canale and Swain's model of communicative competence that would became the theoretical backbone of CLT and acceptance became widespread. Nunan (1998) described it as the most pervasive change to teaching practice over the past twenty years.

According to Richards and Rodgers (2001), the aim of CLT is to promote real-life, authentic communication tasks which are meaningful to the learner and support their learning processes. Holliday (1994) argues there is a strong and weak version of the communicative language approach. The weak communicative approach focuses on language use with an emphasis on student talk time and pair and group work. This approach is underpinned by the belief that communication facilitates learning (Holliday, 1994, p. 170). On the other hand, the strong version focuses on how language works in discourse or how students engage with the text (Holliday, 1994, p. 171). Given the focus on the student and his or her learning in a communicative approach, many teachers mistakenly associate the weak version (e.g. pair and group work) as central to this approach. In classroom contexts in which student numbers are much larger, CLT and a student-centred approach appears contextually inappropriate.

Not surprisingly, many claim that transferring communicative teaching methods to other parts of the world can be problematic because they are not always appropriate to the local context (Ellis, 1996; Holliday, 1994; Kramsch & Sullivan, 1996; Lewis & McCook, 2002; Pham, 2007; Spicer-Escalante & deJonge-Kannan, 2014). On the other hand, however, Pham (2007) argues that, "undoubtedly, CLT originates in the West, but to decide *a priori* that this teaching approach is inappropriate to a certain context is to ignore developments in language teaching" (p. 196). This proves interesting for two reasons. First, there is an assumption that CLT is "Western" and second that to ignore CLT is to ignore language teaching development.

Several empirical studies have been conducted in Vietnam to examine Vietnamese teacher's beliefs about CLT (e.g. Ellis, 1996; Lewis & McCook, 2002; Pham, 2007). In Lewis and McCook's (2002) study, Vietnamese secondary teachers who were attending a workshop on CLT were asked to reflect on teaching practices in Vietnam in the form of journal entries. Though their research may have uncovered some interesting perspectives on how the 14 participating teachers perceived teaching and learning, there was little evidence to suggest that these beliefs were aligned with actual practice.

Similarly, Pham (2007) conducted a study on teachers' beliefs on the use of CLT, in which data is collected from both conversations and classroom observations. Pham selected three female teachers, who had completed graduate degrees from Australian universities, and were teaching at a university in Vietnam. Pham (2007) claims that the teachers in his study embrace CLT but when it comes to applying its principles in practice, they encountered many problems. The teachers did not perceive themselves as being successful, especially with pair and group work. While the teachers expressed their support for CLT practices, their attempt to appropriate them to a Vietnamese classroom offered many contextual and cultural constraints.

2. Washback studies on high-stakes national english examinations

With the growth of communicative language teaching (CLT) in the 1970's, the need for practical measures of language *performance* rose. In contrast with traditional pencil and paper tests, performance testing requires a test taker performance which is observed and judged. McNamara (1996) argues that this shift was in response to two main needs: the need to develop selection procedures (e.g. foreign students studying at English-medium universities) and to align testing with developments in language teaching which had been influenced by theories of communicative competence. Language proficiency, which is currently defined in terms of communicative competence in most high-stakes international language tests (e.g. TOEFL & IELTS), is what is predominately measured in most language tests.

High stakes tests have often been used in language education to change teaching and learning practices as intended by policymakers and test designers (Qi Luxia, 2005). A number of empirical studies have discovered that while test designers have intended for the test to introduce a shift to communicative language teaching practices, they have fallen short of their aim (Andrews, 1994; Chen, 2006; Cheng, 2004, 2005; Wall & Alderson, 1993).

Wall and Alderson (1993) used classroom observation and teacher and student interviews to investigate the washback of the O-Level English examination in Sri Lanka. This test was linked to a textbook series that introduced new ideas in terms of content and methodology (Wall & Alderson, 1993, p. 44). The test and the textbook series was underpinned by a