Australian Journal of Teacher Education

Volume 41 | Issue 7 Article 10

2016

The Washback of the TOEFL iBT in Vietnam

Melissa Barnes
Monash University, Australia, melissa.barnes@monash.edu

The Washback Effects of the TOEFL iBT in Vietnam

Melissa Barnes
Monash University

Abstract: Washback, or the influence of testing on teaching and learning, has received considerable attention in language testing research over the past twenty years. It is widely argued that testing, particularly high-stakes testing, exerts a powerful influence, whether intended or unintended, positive or negative, on both teachers and learners. This article investigates the washback effects of a high-stakes English language proficiency test, the Test of English as a Foreign Language Internet-Based Test (TOEFL iBT), in Vietnam. Vietnam, a developing country whose educational philosophies differ from those underpinning the TOEFL iBT, provided a unique context to explore the test's washback. In the course of this study, four teachers were observed and teaching materials were collected from educational institutions in Vietnam. The study revealed that the TOEFL iBT influenced both what the teachers taught and how they taught but its effects were mediated by the use of test preparation materials.

Introduction

Given that test scores are often used for decision-making purposes, the social consequences of test interpretation and use have received a considerable amount of attention among language testing researchers (e.g. Alderson & Hamp-Lyons, 1996; Andrews, Fullilove, & Wong, 2002; Burrows, 2004; Cheng, 2004; 2005; Hayes & Read, 2004; Messick, 1989; 1996; Shohamy, 2001; Shohamy, Donitsa-Schmidt, & Ferman, 1996; Tsagari, 2011; Wall & Horak, 2006; 2008; 2011; Watanabe, 1996; 2004). Many researchers, realising the social consequences involved in language testing, have attempted to investigate the influence of testing on teaching and learning, or the w*ashback* or *backwash* of a test. Adjectives such as 'intended,' 'unintended,' 'positive' and 'negative' have been used to describe the nature of the washback effect. However, it is argued that empirical evidence to identify the nature of its effects with different tests and in different cultural contexts is still limited (Alderson & Hamp-Lyons, 1996; Bailey, 1999; Zareva, 2005). Thus, the research discussed in this paper aimed to empirically investigate the washback effects of an American English language proficiency test, the Test of English as a Foreign Language Internet-Based Test (TOEFL iBT) in a Vietnamese context in order to contribute to the current understanding of the nature of washback.

Washback Defined

Washback, which is used synonymously with the term 'backwash,' is used in applied linguistics to refer to the influence of testing on teaching and learning. While the concept of

washback had been previously explored in other areas of education, Wall and Alderson's (1993) empirical study and their article exploring the washback phenomenon (Alderson & Wall, 1993) acted as a catalyst for future research in language testing. Wall and Alderson (1993) used classroom observation and teacher and student interviews to investigate the washback of O-level exams, which were used as a national English test in Sri Lanka. Their detailed observations of teacher behaviour, both before and after the implementation, allowed them to distinguish the attitudes and behaviours that could be attributed to the introduction of the test. They found that there were many factors involved, such as teacher ability and knowledge of the test, which added to the complexity of washback. Overall, while the teachers were accepting of the demands of the new test, few of them understood the nature of the test or the methods of the textbooks that they were attempting to teach. This study was the first of many empirical studies to investigate the participants and process of washback (Bailey, 1999) by exploring teachers' and students' attitudes and beliefs about the introduction of the target test, in addition to using classroom observations to further explore teacher and student behaviour.

A large majority of the ensuing empirical research, which has primarily focused on tests that have been modified and improved upon, observed a change in content but little to no change in methodology or teaching methods (Alderson & Wall, 1993; Chen, 2006; Cheng, 2004; 2005; Watanabe, 2004). In Messick's (1996) theoretical review, he argues that, 'A test might influence ***what*** is taught but not ***how*** it is taught... (p. 2).' However, others argue that both content and methodology show evidence of change when a new test is introduced or modified, but this occurs in varying degrees (Alderson & Hamp-Lyons, 1996; Stetcher, Chun, & Barron, 2004).

A major issue in washback studies to date is isolating washback effects from other factors that may be either causing or prohibiting change. Many researchers agree that tests cannot be fully responsible for innovation in teaching and learning, as other factors, such as the teacher variable and test status, play an important role (Alderson & Hamp-Lyons, 1996; Andrews et al., 2002; Burrows, 2004; Cheng, 2004; 2005; Hayes & Read, 2004; Messick, 1996; Shohamy et al., 1996; Spratt, 2005; Tsagari, 2011; Watanabe, 1996; 2004).

Several studies have also highlighted the importance of commercial test preparation materials when determining the washback of a test (Alderson & Hamp-Lyons, 1996; Alderson & Wall, 1993; Cheng, 2004; 2005; Hayes & Read, 2004; Moore, Stroupe, & Mahony, 2012; Tsagari, 2011; Wall & Horak, 2006; 2008; 2011). With a strong international industry dedicated to test preparation textbooks and materials, a test's ability to influence these materials affects both the students and teachers who use them, especially those who are limited in the range of resources available to them. While some researchers (Alderson & Hamp-Lyons, 1996; Hayes & Read, 2004) argue that teachers rely too heavily on test preparation textbooks, others (Cheng, 2005; Hutchinson & Torres, 1994) claim that these textbooks provide the structure and security needed for teachers and learners. Given the importance of materials on 'what' is taught, there is limited empirical evidence as to how helpful these materials are in preparing for the target test.

Target Test: TOEFL iBT

In May 1961, a conference was held by the National Association of Foreign Student Advisers (NAFSA) and the Institute of International Education (IIE) in Washington which aimed to establish a battery of language tests to assess English language proficiency skills of non-native speakers who desired to study at universities and colleges in the United States. This resulted in the creation of the TOEFL, the Test of English as Foreign Language, in 1962.

Situated within a psychometric testing paradigm, the test was objective, machine-scored, cost-effective and profitable, and secure and efficient, in contrast with more subjective, integrated and human testing approaches (Spolsky, 1995, p. 217). The test originally consisted of five sections: structure, listening, vocabulary, reading and writing. In 1976, the Paper-Based Test (PBT), was introduced with three subtests assessing listening, writing, and reading skills and a structure (grammar) subtest was added in subsequent years.

In 1995, a call for change was initiated by various constituencies, consisting of score users and second language teaching and testing experts, who believed that the test should reflect communicative competence models, include more constructed-response tasks, directly measure writing and speaking, integrate language skills, and measure a student's ability to communicate in an academic setting (Educational Testing Service, 2007). The new TOEFL test was introduced in two phases. the Computer-Based Test (CBT) was introduced as an interim test in 1998 and consisted of the previous TOEFL test design with some enhancements such as a computer-mediated format. Seven years later, in 2005, the second phase of the TOEFL project was released with the rollout of the TOEFL iBT.

The TOEFL iBT differs from previous formats (Paper Based Test [PBT] and Computer Based Test [CBT]) in that it focuses on all four macro language skills (speaking, listening, reading and writing) and academic communication, and is underpinned by an integrated approach. TOEFL iBT is the first TOEFL test to include a speaking section, in which structure (grammar) is assessed through the speaking and writing sections rather than as a separate subtest. Although the TOEFL iBT shares a computer-mediated format with its predecessor the CBT, its introduction of a semi-direct speaking sub-test requires test-takers to speak into a microphone attached to their headset so that a digital file can be recorded. Another main feature of the iBT is its focus on academic communication. For example, students listen to longer conversations and lectures set in an academic context (i.e. a student asking a librarian questions, a Geography lecture, etc.) and are encouraged to take notes, which was not allowed in previous formats. Integrated tasks, in which students gather information from a variety of sources and respond with a written or spoken response, reflect authentic academic communication and skills needed to be successful in an academic setting. The writing and speaking sections consist of both independent tasks (based on test taker's opinion and background knowledge) and integrated tasks (based on written and spoken texts provided within the test).

While washback research on the TOEFL iBT is still quite limited, Wall and Horak's (2006; 2008; 2011) 5-year longitudinal study, which explored the influence of the iBT on teaching and learning in Central and Eastern Europe, provides a significant contribution to understanding the washback effects of the TOEFL iBT. Phase 1 (Wall & Horak, 2006) acted as a baseline study, or an antecedent, with the aim being to describe what TOEFL preparation courses in Central and Eastern Europe looked like before the introduction of the TOEFL iBT. They observed 10 TOEFL classes in six countries and interviewed 10 teachers, 21 students and nine directors. Wall and Horak (2008) found that teachers did not express as much negativity about teaching the TOEFL (Computer-based) as was reported by Alderson and Hamp-Lyons (1996), who explored the washback of the TOEFL (Paper-based) on TOEFL preparation courses in the United States. However, the reliance on test preparation textbooks was reported in both studies.

Phase 2 focused on six of the teachers who Wall and Horak (2008) had been working with in Phase 1. Their aim was to gather qualitative data regarding the teachers' awareness, preparedness and attitudes toward the iBT just before it was released. Their findings from Phase 1 and 2 exposed the importance of test preparation coursebooks (textbooks) as they were 'at the heart' of the courses they were examining (Wall & Horak, 2006, p. 78). Therefore, Phase 3 investigated the content and methodology associated with the textbooks

(for both the Computer-Based Test and Internet-Based Test) used by four of the six teachers from Phase 2.

The final stage (Phase 4) observed and interviewed three of the same teachers one year after the introduction of the TOEFL iBT to discover what their classrooms looked like and to determine what aspects of teaching changed. Wall and Horak's (2011) work brought to light several aspects of the TOEFL iBT's influence on test preparation courses:

1. Content changed considerably due to the changed format and tasks dictated by the new textbooks and, while it may not have been to the same extent, there was evidence that methodology changed slightly due to the introduction of a speaking component. Two of the participating teachers increased time allocated for speaking in their classes from 0 to 35%, while the other from 5 to 20%.
2. The textbooks designed for and used in preparation courses were very powerful as they not only dictated what was taught and how students learnt but also teacher behaviour. Teachers were unsure of what the TOEFL iBT would require of them in regards to teaching, so the textbooks provided them with security in knowing what and how to teach the content. '…our impression at the time of the investigation was that they would not have the desire, the need, or the time to stop depending on published materials in the future' (Wall and Horak, 2011, p. 133).
3. Communication between the test designers and teachers and students and between testing agencies and authors and publishers who design the textbooks was extremely important.

While Wall and Horak's (2011) work has been significant in understanding the washback effects of the TOEFL iBT, particularly teacher behaviour and attitudes before and after the rollout of the test, it is limited to its effect on test preparation courses in Europe. The following study contributes to the current understanding of washback by looking at the TOEFL's influence on English language programs in a Vietnamese context.

Methods

In 2009, research was conducted to investigate the washback of the TOEFL iBT in Vietnam. Vietnam was chosen as the context for this study because of its growing market for study abroad programs in English-speaking countries and history of more traditional teaching methods. Vietnam is a country whose approach to education has been shaped by its political relationships and therefore provides an interesting backdrop to study the washback effects of a test which is underpinned by language learning philosophies very different from its own.

This case study aimed to provide a holistic depiction of washback by comparing several variables. Materials were collected from five TOEFL iBT classrooms and four teachers (two native and two non-native speakers of English) were observed in both TOEFL iBT preparation and general English courses in order to isolate the test's influence on what is taught and how it is taught in test preparation courses. In order to investigate the washback effects of the TOEFL iBT, two research questions shaped the focus of this study:

- What are the effects of the introduction of the TOEFL iBT on what is taught (content)?
- What are the effects on how it is taught (teaching methodology or pedagogy)?

For the purpose of this study, content refers to authentic and commercial teaching materials and textbooks utilized within the classroom. Teaching methodology or pedagogy, on the other hand, refers to the teaching methods and activities employed in the classroom, which are underpinned by how teachers believe their students learn. While washback offers a variety of research angles to pursue, this case study focuses primarily on how the TOEFL

iBT influences pedagogy, or more specifically teaching and teachers, rather than also investigating its influence on learning and students. Given the study's focus on teachers, it was important to find participants that could provide a range of perspectives and represent a variety of classroom settings.

Participants

English language education in Vietnam, Hanoi in particular, is represented by the private and public sectors; therefore, it was important to not only select institutions that were from both of these sectors, but to also find institutions that taught and promoted the TOEFL iBT.

Two language centres were chosen for this study, an American language centre, which will be referred to as AL, and a Vietnamese language centre, which will be referred to as VL. Both language centres taught both general English courses in addition to test preparation courses such as TOEFL iBT. The national university chosen for this study, which will be referred to as NU, consists of several campuses, which include discipline-specific branches and centres. Key contacts were established at each institution and approval granted. These contacts, then, provided names and contact information of teachers who were currently teaching both TOEFL iBT and general English language courses. While six teachers expressed interest, four of the six were asked to participate in order to have classroom contexts from both the public and private sectors and with non-native and native-English speaking teachers. One of the NU teachers also taught private classes (PC) out of his home, providing another window into the private sector. The teachers ranged from 24 years of age to 60. Two were Vietnamese, one American and the other British. Three of the four teachers had Teaching English as a Foreign Language (TEFL) or Teaching English to Speakers of Other Languages (TESOL) qualifications; however, one of these three had completed his qualification online as he was unable to take a regular course without a university degree. The only female participant had recently graduated top of her class at NU and was chosen to stay at the university as a lecturer. Teaching experience among the teachers ranged from 15 years to 5 months. Table A provides a brief profile of the four participating teachers, illustrating the wide range of backgrounds and educational and professional experiences present among them. All participating teachers were given pseudonyms to protect their identities.

	Tuan	**Mike**	**David**	**Ly**
Gender	Male	Male	Male	Female
Nationality	Vietnamese	American	British	Vietnamese
First language	Vietnamese	English	English	Vietnamese
Academic qualifications	Bachelor in TESOL—Vietnam Masters in Linguistics—Australia	Bachelor in Economics—USA TEFL certification--Thailand	High school diploma—UK Online TEFL certification	Bachelors in Education—Vietnam
Years teaching English	15	5 months	3	2
Institution of employment	NU Self--PC	AL NU	AL	VL NU

Table A: Background of Participating Teachers

While student data was not collected in order to keep the scope of this study manageable, student consent was needed to observe the participating classrooms. Consent forms were distributed and collected by the researcher on arrival in Vietnam and translation was provided. The age of students ranged from 18-45 years and the classes ranged in size from 10-45 students in one class.

Procedure

In order to investigate the washback effects of the TOEFL iBT, data were collected on the content (i.e. curriculum and teaching materials) and pedagogy (i.e. teaching methods and activities) through collecting classroom materials and observing classrooms. The data for this study were collected from April 9 to May 15, 2009 in Hanoi, Vietnam.

Classroom Materials

The first research question, 'What are the effects on what is taught?' was addressed by collecting teaching materials from the TOEFL iBT classes. Teaching materials, such as student worksheets, textbooks, CDs and DVDs, PowerPoint presentations and other supplementary materials, were collected for further analysis. In addition, field notes were taken to document the resources, such as computers, whiteboards, and audio equipment, that were available to teachers and students in the classrooms observed. Given that the content of a course is often realised through the materials and resources that are employed, the collection and analysis of teaching materials provided an opportunity to not only explore what was being taught in the classes observed but to discover the alignment between what was being taught and what was being introduced in the new format of the TOEFL. The materials were collected by the researcher during the classroom observations and then divided into constituent 'tasks' or activities, which were analysed in further detail (see Table B). The tasks were analysed using Littlejohn's (1998) framework, which was specifically designed to analyse English language materials by dividing the materials into constituent tasks for an in-depth investigation of the *process*, *classroom participation* and *content*.

Classroom Observations

In order to answer the second research question, 'What are the effects on how the TOEFL iBT is taught?' teachers representing the private and public sectors and native and non-native English speakers were observed in an iBT preparation class and a general English language class. Given that the rollout of the TOEFL iBT had already begun at the time of the study and a baseline study could not be conducted, the general English classes acted as a point for comparison. These classes ranged from 1.5 to 3 hours per lesson. While three of the teachers were observed twice, once in an iBT course and once in a general English course, the fourth teacher was observed teaching in an iBT course at both a language centre and a national university and a general English course in the national university. The rationale was to distinguish between a teacher's individual teaching style and the methods they use to adapt to their teaching context. Due to ethical considerations, the lessons were not audio recorded or video recorded.

The Communicative Orientation of Language Teaching Observation Scheme (COLT [Spada & Frohlich, 1995]) was employed to analyse classroom events by identifying methods used and recording the various applications of content materials in real time. While COLT

consists of Part A (activity level) and Part B (exchange level with a focus on linguistic features used within the classroom), only Part A was relevant to the research design adopted for this study.

Part A of the COLT requires the observer to make detailed notes in real time on the activities and episodes that occur during the course of the lesson. Classroom activities, or events, are viewed as the unit of analysis, rather than a time period, and are analysed by a set of themes or codes established by Spada and Frohlich (1995). The coding conventions are used to identify five main components: *time, participant organisation, content, student modality* and *materials*. In addition to these five components, an analysis of *activity type* as a percentage of class time is employed (see also Cheng, 2005) as an adjunct to the *participant organisation* and *content* components.

Findings and Discussion

Through the analysis of teaching materials and teaching methods, this study reveals the complexity of washback and the degree to which the TOEFL iBT has influenced classroom content and teaching methods.

The Influence of TOEFL iBT on Classroom Materials

This study found that teachers heavily relied on TOEFL iBT-specific materials for instruction and classroom activities. All four institutions used a set textbook or a collection of pages from TOEFL iBT textbooks in their courses. The Vietnamese Language centre (VL) and Tuan's private class (PC) were based around a collection of photocopied tasks from commercial TOEFL iBT textbooks, which were given to the students as a bound set (VL) or given out separately to students at each class (PC). The textbooks or collections used by all of the participating institutions were from commercial TOEFL iBT textbooks published primarily in the years 2006 and 2007 (See Appendix A for a reference list). It is important to note that none of the textbooks employed by the teachers were from Educational Testing Service (ETS), the governing body of the TOEFL iBT, but were sourced from external publishing companies. These textbooks followed a similar organisational pattern; the chapters were organised into subtests or macro language skills, such as listening, speaking, writing and reading, and then each subset was divided into academic skills required for that subset. These academic skills were identified by textbook designers as necessary skills needed for taking each subtest. Only one of the textbooks, which was used as a supplementary text, had an accompanying teacher's manual.

In collecting data from the five TOEFL iBT classes observed, sixteen separate tasks or activities were identified and collected for further analysis. It was found that the majority of the materials collected in the class observations were not created or sourced by the teachers or students, but came directly from the classroom textbook. Only two of the sixteen tasks were from non-TOEFL iBT related sources. Both of these tasks were used by the same teacher in the same lesson. David began his TOEFL iBT class with an integrated speaking task from his TOEFL iBT textbook, but then moved onto two non-TOEFL iBT material sources, a list of discussion questions and a reading passage he had taken from an online website. He was the only teacher to use non-TOEFL iBT materials, which suggests that TOEFL iBT textbooks play an important role in determining what content is taught in the classroom.

Ten Tasks Analysed

Of the sixteen tasks identified and collected from the classroom observations, ten tasks were chosen and analysed, using a framework by Littlejohn (1998). These ten tasks were chosen to represent the five TOEFL iBT lessons observed, with two tasks from each participating teacher's lesson (two of the four teachers only used two tasks in the entirety of their lesson). When determining which task to analyse when a teacher employed more than two tasks in one lesson, the task was chosen based on fairly representing macro language skills (e.g. speaking, listening, reading, and writing). These tasks were then analysed by exploring the processes, participation and content.

1. The **process** students and teachers must go through in the task.
   Guiding question: ***What is the student expected to do?***
2. The **classroom participation**, which defines who students are to work with, if anyone.
   Guiding question: ***Who with?***
3. The **content** the students are to focus on.
   Guiding question: ***With what content?***

By way of answering Littlejohn's (1998) first two guiding questions regarding process and participation, the majority of tasks analysed expected students to *respond* to questions *individually*. Table B provides a summary of the ten tasks analysed, which details the macro skill in focus, a brief description of the task (in regards to the language skill(s) being developed), the type of input given to the students, the output that was expected from the students and the source of the material. Appendix A provides the reference information for each text as indicated in the 'source' column. Of the 10 tasks, Task 6 (Non-TOEFL iBT: Conversation questions on age), Task 1 (pronunciation of numbers and symbols), Task 5 (integrated speaking task) and Task 10 (integrated speaking task), required students to respond to the class or in pairs and groups. While the tasks had no instructions indicating with whom students should participate, the task required a spoken response in which some sort of interaction or feedback was assumed. All of the tasks that required an oral response did not specify the intended classroom participation or with whom the students should be interacting. For example, Task 5, an integrated speaking task, states, 'The woman [in the previous listening exercise] expresses her opinion of the Career Services Center. State her opinion and explain the reasons she gives for holding that opinion.' The task describes what the speaker is to do but not how or who with. Given that many textbooks are designed for both classroom instruction ***and*** as an individual resource, they do not provide specific teacher guidelines on how these tasks are to be implemented as classroom practice.

	Macro Skill	Description	Input	Output	Source
Task 1	Speaking: Pronunciation	Numbers & symbols	Written words/sentences	Oral words/sentences	TOEFL iBT vocabulary book (Text F*)
Task 2	Listening Reading Writing	Integrated writing task	Spoken & written discourse	Written discourse	TOEFL iBT textbook (Text B*)
Task 3	Reading	Identifying topics and paraphrasing	Written discourse	Written words/sentences & answering multiple choice	TOEFL iBT textbook (Text D*)

Task 4	Listening	Understanding main ideas & organization	Spoken discourse	Written words/sentences & answering multiple choice	TOEFL iBT textbook (Text D*)
Task 5	Speaking	Conversation questions about age	Spoken & written words/sentence	Oral discourse	TOEFL iBT textbook (Text D*)
Task 6	Speaking Listening Reading	Integrated speaking task	Spoken & written discourse	Oral discourse	Non-TOEFL material (Internet A*)
Task 7	Listening	Summarizing a process	Oral discourse	Multiple choice	TOEFL iBT textbook (Text C*)
Task 8	Listening	Placing steps in a sequence	Oral discourse	Multiple choice	TOEFL iBT textbook (Text A*)
Task 9	Listening	Understanding the details	Oral discourse	Multiple choice	TOEFL iBT textbook (Text E*)
Task 10	Speaking	Integrated speaking task	Oral words/sent. & written words/sent.	Oral discourse	TOEFL iBT (Text C*) textbook
***Refer to Appendix A for reference information**					

Table B: TOEFL iBT tasks

While the majority of tasks analysed shared the same process (students were expected to respond) and classroom participation (students answered questions individually), the content among the different tasks varied. First, in analysing the type of input that was provided to students, oral discourse of 50 words or more was used to elicit responses from students in six out of the 10 tasks. Overall, written and spoken texts were preferred over words, phrases and sentences, with four tasks utilising oral discourse, one task written discourse and two tasks incorporating both oral and written discourse. This suggests that the TOEFL's aim to incorporate tasks that require students to gather information from a variety of written and spoken sources in order to respond to a task is reflected in the materials analysed.

Similarly, the output expected from the students in these tasks reflects the output expected of students taking the TOEFL iBT test. Half of the tasks selected required students to respond by answering multiple-choice questions, with three of these tasks requiring only multiple choice answers and two requiring both multiple choice and short answers. The other half, however, students were asked to provide an oral or written response, with all but one at the discourse level. Only one of these tasks asked students to produce a written text, while the others focused on oral responses. Overall, the majority of the tasks required students to respond through test-like multiple choice and short answers on listening and reading subtests or with an oral response on integrated tasks.

In conclusion, the type of input that the students are receiving includes longer spoken and written texts as opposed to words, phrases and sentences (e.g. students must read an article and listen to a discussion before providing a written or spoken response) and the majority of the tasks require students to respond by answering multiple-choice and short answer (e.g. tasks in the listening and reading subtests) and written and spoken discourse (e.g. individual and integrated tasks). Therefore, it appears the introduction of the TOEFL

iBT has influenced the content found in the TOEFL iBT textbooks used in the lessons observed due to the similarity of the input and output found on the test and the test-specific materials. However, this influence goes as far as what ***can be*** taught in the classroom and not what actually happens in the classroom. Given the lack of teacher instructions and details pertaining to classroom participation, there is the potential for a discrepancy between what the writers of the tasks intended and how they are employed in the classroom. How these ten tasks were realised in the classroom will be discussed below.

The Influence of the TOEFL iBT on Teaching Methods

The findings from the analysis of the teaching materials suggest that TOEFL iBT textbooks, which seek to mirror the tasks found on the test, play an important role in ***what*** is taught. However, this section attempts to explore ***how*** these materials were actually taught. The data collected from the *Communicative Orientation of Language Teaching Observation Scheme* (Spada & Frohlich, 1995) during the classroom observations provided insight into participant organisation (similar to classroom participation) and activity types.

Patterns of participant organisation in the COLT are described by Spada and Frohlich (1995) in terms of: 1) is the teacher working with the whole class, 2) are students working in groups, or 3) are they working individually? Patterns of participant organisation found in both the iBT preparation and general English classes are reported in Figure A below, which shows the participant organisation patterns in the lessons observed in this study as a percentage of total class time. The whole class represents interactions between the teacher and student or the entire class (T to S/C), and student to students or class (S to S/C). Examples of student to student or class interaction (S to S/C) include oral presentations and class discussions, which are initiated and controlled by the students.

As can be seen from Figure A, all but one (Mike AL) of the five TOEFL iBT classes observed spent more than half of the class time on teacher and student or class interaction (T to S/C), indicating a focus on teacher instruction. The general English classes were not much different, with two of the four general English classes spending more than half of the class time on teacher instruction. With the exception of Mike's TOEFL iBT lesson at AL, all of the TOEFL iBT classes had a higher percentage of teacher instruction than in their general English classes. Interestingly, however, although Mike and David incorporated group work in their general English classes, Mike did not use group activities in either of his TOEFL iBT lessons and David's use of group work was a result of using the only non-TOEFL source (conversation questions). Also, the majority of class time in all of the TOEFL iBT lessons was spent on teacher instruction and individual student practice. Ly and Tuan spent all of their class time on teacher instruction and individual student practice, while Mike spent a combined 88% (AL) and 89% (NU) and David spent 65%, due to his use of the non-TOEFL related conversation questions. Overall, the TOEFL iBT classes focused primarily on teacher instruction and individual student practice, which aligns with the results from the materials analysis in that students were expected to respond individually to the test tasks found in TOEFL iBT textbooks.

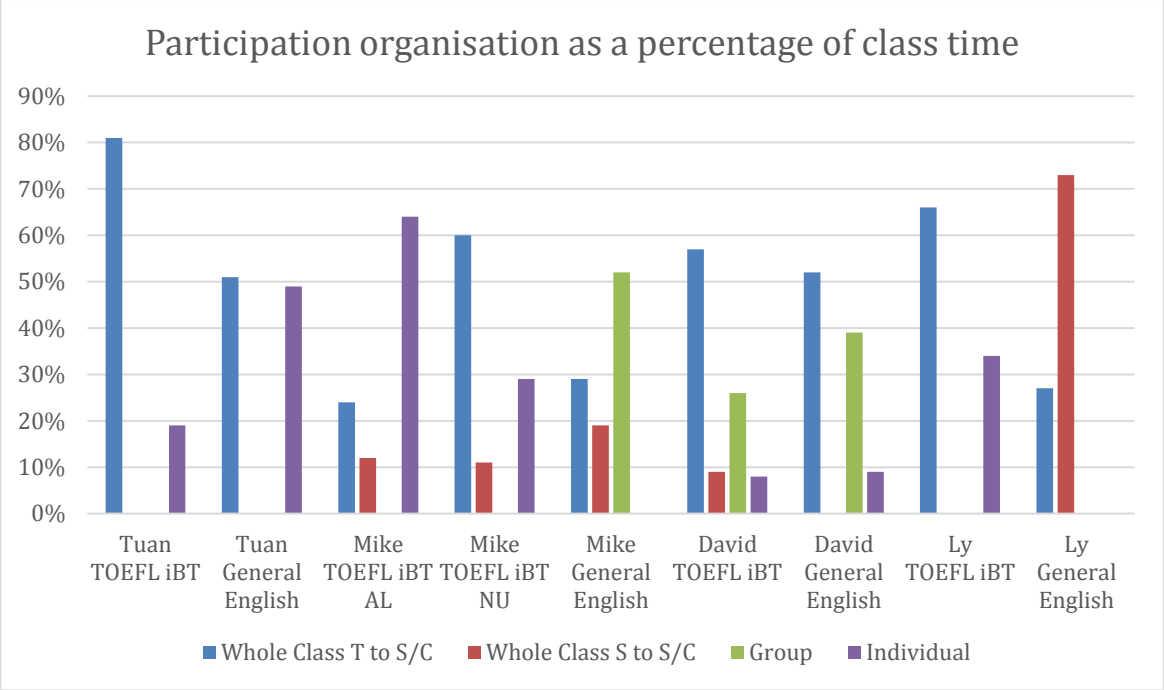

Figure A: Participant organisation as a percentage of class time

By investigating the type of activities observed in this study, a clearer picture develops of how teaching and learning is realised in the classes observed. Patterns of participant interaction are further explored by identifying teacher activities, class activities and student activities. These activities are then classified by the primary purpose or nature of the activity (i.e. lecturing, explaining, guiding). In Figure B below, the activities are reported as a percentage of class time so as to make judgements about what activity types are given priority in the classroom. Overall, the table reinforces the predominant role of the teacher and the role of individual work in TOEFL iBT preparation classes.

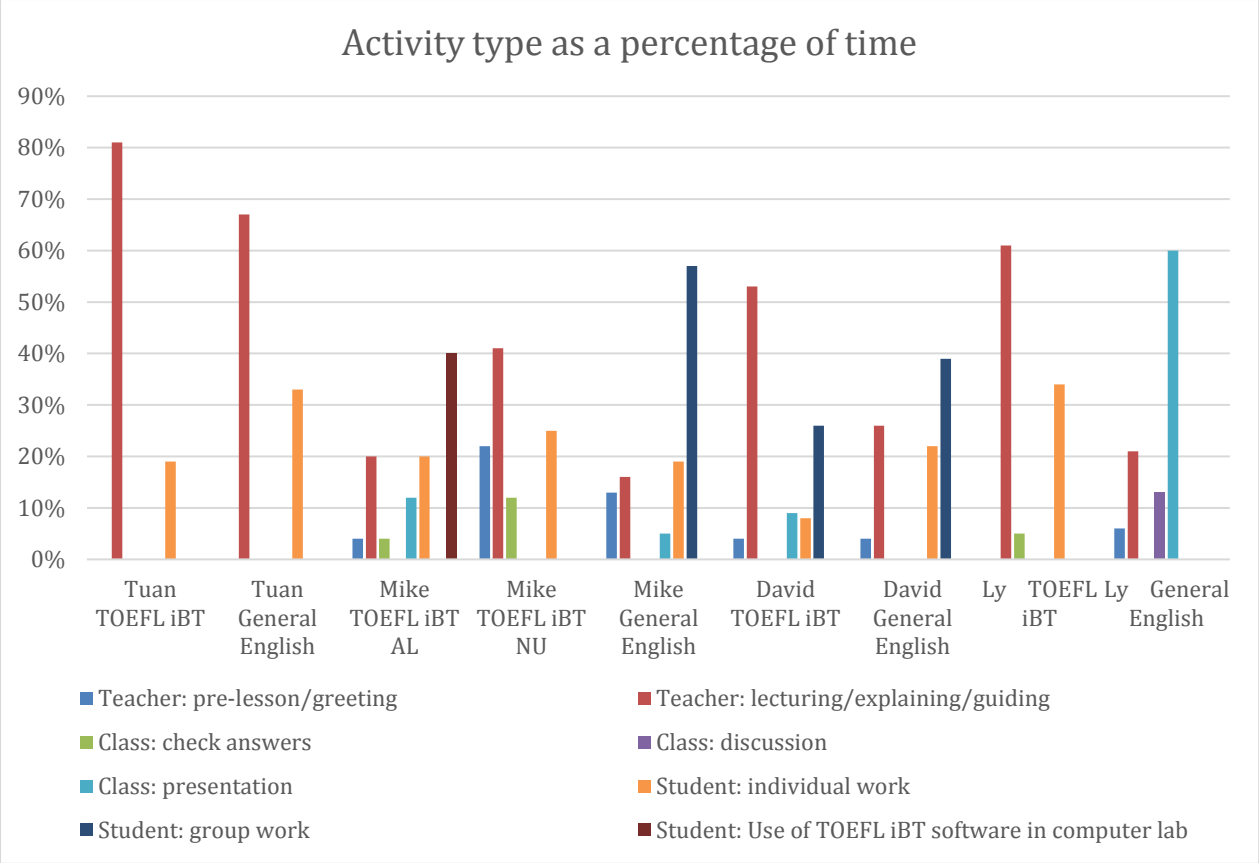

Figure B: Activity types as a percentage of class time

As was reported in Figure B, teacher instruction, or teacher activities more broadly, occupied more than half of the total class time in all of the TOEFL iBT preparation classes, with the exception of Mike's AL class which spent only 24% of total class time on teacher activities. Given that Mike's AL class was the only class, which had scheduled time in the computer lab during the lessons observed, which took 40% of class time, the percentage of teacher activities in his AL class was replaced by student practice time. It should also be noted that the computer lab session was treated as its own activity type and not as an independent student activity in order to distinguish between individual practice with textbooks and with software that more accurately reflects the test format and mode. Overall, four out of the five TOEFL iBT preparation courses had a higher percentage of teacher activities than their corresponding general English class. It was only Mike's AL course that had a slightly lower percentage (24%) than his general English course (29%), which again, may have been a result of having spent a large part of the lesson in the computer lab in which students were working independently and teacher activities were kept to a minimum.

According to the data collected from the COLT, students received few opportunities to speak to one another in English (pair and group work) or as part of a class discussion. While all of the teachers explained or gave the answers to the practice TOEFL iBT tasks, Mike and Ly involved students by calling on them to answer the questions instead of just giving them the answers. However, the time spent was limited, with Mike spending 4% of class time in his AL lesson and 12% with his NU class and Ly spending 5%. Tuan and David tended to walk the students through an exercise, such as a listening text, and gave the answers as they went. However, another way in which speaking opportunities were provided to students in class was observed in Mike and David's classes, in which they attempted to engage students in speaking activities by providing opportunities for students to present their constructed responses to the entire class, devoting respectively 12% and 9% of total class

time to this. While Mike gave every student in the class an opportunity to respond in front of the class and receive feedback, David only chose one student. Though David's feedback was thorough, only one of his students received a chance to respond and receive feedback on that particular day. As discussed before, group and pair work were also limited in the classrooms observed. In contrast to the limited amount of class activities in the other nine lessons observed, Ly's general English class dedicated 60% of class time to class activities, as oral presentations were an important feature of her speaking course.

Student activities in the TOEFL iBT preparation courses averaged around 34% of total class time, while the general English classes averaged about 43%. Overall, general English classes provided more student practice than the TOEFL iBT preparation courses, which may not come as a surprise due to the high number of teacher activities. However, the dominant use of textbooks, which primarily include practice exercises, might lead one to believe that there would be more practice in the TOEFL iBT preparation courses, as students are required to work through the practice exercises. However, in the TOEFL iBT preparation lessons observed, teacher instruction tends to occupy more class time than student activities.

Overall, while most of the TOEFL iBT preparation courses observed spent the majority of class time on teacher instruction or explanations, in addition to incorporating individual practice rather than pair and group work, the general English classes tended to be more student focused with a large amount of time spent on student activities. The findings of this study suggest that while the teachers taught the content found in the textbooks, which reflect the test tasks on the TOEFL iBT, they allowed the nature of these test-like tasks to ***limit*** their approach to teaching. While there were opportunities for students to interact with one another allowing for more individual feedback and the co-construction of language, seldom were these opportunities given. Therefore, the format of the TOEFL iBT textbooks may encourage teachers to adapt an approach to teaching that focuses primarily on instruction and individual response; albeit this type of input and output may be appropriate for test-taking it may not be appropriate for the classroom.

An underlying principle of the TOEFL iBT textbooks used in this study, and thus the test preparation courses that employ them, is that the language skills that are tested are the skills that should be taught. TOEFL iBT preparation courses, then, focus primarily on the skills needed for the test and the not the process of ***acquiring*** language skills or the pedagogy employed to support this process. Widdowson (1981) makes a distinction between 'goal-oriented courses' and 'process-oriented courses' or in other words, what students need to do with language once they have learned it and what students must do to acquire language. By these definitions, the TOEFL iBT preparation courses observed reflect a 'goal-oriented' course as the content of these courses is focused on the end goal (the test) rather than on the process of acquiring language. This is problematic as teachers may then approach teaching test preparation courses as completing a textbook that acts as a 'skills checklist,' as suggested by the data from this study, instead of approaching teaching and learning more holistically.

Conclusion

The TOEFL iBT is a powerful, international language test and the findings from this study suggest that it has the ability to influence ***what*** is taught and to some degree ***how*** teachers teach in English language programs in Vietnam, particularly in TOEFL iBT preparation courses.

The results suggest, in alignment with other washback studies (Alderson & Hamp-Lyons, 1996; Alderson & Wall, 1993; Cheng, 2004; 2005; Wall & Horak, 2008; 2011), that

commercial test preparation materials are central to the teaching and learning in test preparation courses. While Wall and Horak (2011, p. 49) also came to the conclusion that TOEFL iBT textbooks were 'at the heart of each teacher's lesson,' this study provides empirical data on how these materials were realised in the classroom. These textbooks were very instrumental in determining what was taught in the TOEFL iBT preparation courses observed, particularly the inclusion of speaking and integrated tasks in classroom activities.

In addition, while a number of empirical washback studies observed a change in content but little to no change in pedagogy (Alderson & Wall, 1993; Chen, 2006; Cheng, 2004; 2005; Watanabe, 2004), the findings from this study suggest that teaching methods were slightly influenced due to their reliance on TOEFL iBT specific materials. While there were not significant differences or a uniform pattern between participant organisation and activity types in the TOEFL iBT and general English classes observed, the TOEFL iBT classes were very similar in that the overwhelming majority of class time was spent on explaining the tasks or activities in these textbooks (teacher instruction) and then allowing the students to complete the related exercises (individual student practice). It can be argued that TOEFL iBT textbooks have the ability to influence teacher behaviour and methods in the classroom because teachers choose to follow these materials closely to ensure they teach all the necessary language skills that are assessed on the test; therefore, limiting preferred teaching styles and methods. This was evidenced in Mike and David's classes, as they incorporated more opportunities for students to interact and speak to one another in their general English classes than they did in their TOEFL iBT classes. These findings align with those of Wall and Horak (2011) in that TOEFL iBT textbooks have the ability to influence teacher behaviour and therefore have an influence on how teachers teach. Given that TOEFL-iBT preparation classes are often highly structured and goal-oriented, teachers and their preferred teaching methods are often overshadowed.

This study argues that the introduction of the TOEFL iBT has had an influence on English language classes in Vietnam but its influence is mediated by the influence of TOEFL iBT textbooks on both what is taught and how it is taught in the classroom. The implications of these findings are methodological, pedagogical and theoretical in nature. First, given the important role of textbooks in the classroom, there is a need to reconsider our methods in analysing the accuracy and adequacy of the TOEFL iBT materials on the market and explore how these materials impact teaching and student learning. Questions need to be asked in regards to how well test preparation textbooks accurately represent what is on the test, how well they support student learning and finally how well they encourage positive teaching and learning practices. Secondly, in order to support and guide teaching pedagogy, particularly for teachers who are inexperienced or lack confidence in teaching the material, there is a need for textbooks to be specifically designed for classroom use and have accompanying teacher manuals and guides. While these do in fact exist, their availability was limited in Vietnam at the time of this study. Finally, the theoretical implications of this study include the need for better communication between the test designers and textbook publishers and authors (Wall & Horak, 2011) to ensure that the TOEFL iBT textbooks adequately reflect the skills underpinning the TOEFL iBT and adhere to a communicative and integrated approach. While a test could be underpinned by a communicative and integrated approach and be validated by research, if materials are not specifically designed for classroom instruction, test preparation courses may not be able to truly reflect the communicative constructs that the test designers had in mind.

While this study shed light on how TOEFL iBT materials influenced the content and pedagogy of the test preparation courses, it lacked student and teacher perspectives and attitudes towards teaching, learning and testing. Not only would this have allowed for a richer understanding of the influence of the TOEFL iBT but it would have allowed for more data on

the quality of student learning, an area that needs to be further researched. In addition, one of the biggest weaknesses of this study was that it did not observe more teachers over a longer period of time, particularly before the TOEFL iBT was rolled out. Observing teachers before and after the introduction of the TOEFL iBT would have provided data on the content and teaching methods used in the previous TOEFL tests, allowing for a better understanding of the extent of the TOEFL iBT's influence.

References

Alderson, J. C. & Wall, D. (1993). Does washback exist? Applied Linguistics, 14, 115-129. http://dx.doi.org/10.1093/applin/14.2.115

Alderson, J.C. & Hamp-Lyons, L. (1996). TOEFL preparation courses: A study of washback. Language Testing, 13(3), 280-297. http://dx.doi.org/10.1177/026553229601300304

Andrews, S., Fullilove, J., & Wong, Y. (2002). Targeting washback. System, 30, 207-223. http://dx.doi.org/10.1016/S0346-251X(02)00005-2

Bailey, K. (1999). Washback in language testing (TOEFL Monograph Series 15). Princeton, NJ: Educational Testing Service.

Burrows, C. (2004). Washback in classroom-based assessment: A study of the washback effect in the Australian Adult Migrant English Program. In Cheng, L, Watanabe, Y., & Curtis, A. (eds.), Washback in language testing (pp. 113-128). Mahwah, NJ: Lawrence Erlbaum Associates.

Chen, L. (2006). Washback effects on curriculum innovation. Academic Exchange, 206-210.

Cheng, L. (2004). The washback effect of a public examination change on teachers' perceptions towards their classroom teaching. In Cheng, L, Watanabe, Y., & Curtis, A. (Eds.), Washback in language testing (pp. 147-170). Mahwah, NJ: Lawrence Erlbaum Associates.

Cheng, L. (2005). Changing language teaching through language testing: A washback study. Cambridge: Cambridge University Press.

Hayes, B., & Read, J. (2004). IELTS test preparation in New Zealand: Preparing students for the IELTS academic module. In Cheng, L, Watanabe, Y., & Curtis, A. (Eds.), *Washback in language testing* (pp. 97-111). Mahwah, NJ: Lawrence Erlbaum Associates.

Hutchinson, T., & Torres, E. (1994). The textbook as agent of change. *ELT Journal, 48*(4), 315-328. http://dx.doi.org/10.1093/elt/48.4.315

Littlejohn, A. (1998). The analysis of language teaching materials: Inside the Trojan Horse. In B. Tomlinson (Ed.), Materials development in language teaching (pp. 190-216), Cambridge: Cambridge University.

Messick, S. (1989). Perspectives on the evolution and future of educational measurement. In Linn, R. (ed.). Educational Measurement, third edition (pp. 13-104). New York: Macmillan.

Messick, S. (1996). Validity and washback in language testing. Princeton, NJ: Educational Testing.

Moore, S., Stroupe, R., & Mahony, P. (2012). Perceptions of IELTS in Cambodia: a case study of test impact in a small developing country. In J. Osbourne (Ed.), *IELTS Research Report 13*(6): 1-109. Melbourne: IDP IELTS Australia.

Shohamy, E. (2001). The power of tests: A critical perspective of the uses of language tests. Harlow, England: Longman.

Shohamy, E., Donitsa-Schmidt, S., & Ferman, I. (1996). Test impact revisited: Washback effect over time. Language Testing, 13(3), 298-317. http://dx.doi.org/10.1177/026553229601300305

Spada, N. & Frohlich, M. (1995). COLT: Communicative orientation of language testing observation scheme coding conventions and applications. Sydney: National Centre for English Language Teaching and Researching Macquarie University.

Spolsky, B. (1995). *Measured words*. Oxford: Oxford University Press.

Spratt, M. (2005). Washback and the classroom: The implications for teaching and learning of studies of washback from exams. *Language Testing Research*, 9(1), p. 5-29. http://dx.doi.org/10.1191/1362168805lr152oa

Stetcher, B., Chun, T., & Barron, S. (2004). The effects of assessment-driven reform on the teaching of writing in Washington State. In Cheng, L, Watanabe, Y., & Curtis, A. (Eds.), *Washback in language testing* (pp. 53-71). Mahwah, NJ: Lawrence Erlbaum Associates.

Tsagari, D. (2011). Washback of a high-stakes English exam on teachers' perceptions and practices. In Kitis, E., N. Lavidas, N. Topintzi & T. Tsangalidis (Eds.), Selected papers from the *19th International Symposium on Theoretical and Applied Linguistics (*pp. 431-445*.)*, Department of Theoretical and Applied Linguistics, Aristotle University of Thessaloniki, School of English, Aristotle University of Thessaloniki, Greece.

Wall, D. and Alderson. J. (1993). Examining washback: The Sri Lankan impact study. *Language Testing, 10*, 41-69. http://dx.doi.org/10.1177/026553229301000103

Wall, D. and Horak, T. (2006). The impact of changes in the TOEFL examination on teaching and learning in Central and Eastern Europe. Phase 1: The baseline study (TOEFL Monograph 34). Princeton, NJ: Educational Testing Service.

Wall, D. and Horak, T. (2008). The impact of changes in the TOEFL examination on teaching and learning in Central and Eastern Europe. Phase 2: Coping with change (TOEFL iBT-05). Princeton, NJ: Educational Testing Service.

Wall, D. and Horak, T. (2011). The impact of changes in the TOEFL exam on teaching in a sample of countries in Europe and learning in Central and Eastern Europe. Phase 3, The role of the coursebook. Phase 4, Describing change. (TOEFL iBT-17). Princeton, NJ: Educational Testing Service.

Watanabe, Y. (1996). Does grammar translation come from the entrance examination? Preliminary findings from classroom-based research. Language Testing, 13(3), 318-333. http://dx.doi.org/10.1177/026553229601300306

Watanabe, Y. (2004). Teacher factors mediating washback. In Cheng, L, Watanabe, Y., & Curtis, A. (Eds.), Washback in language testing (pp. 129-146). Mahwah, NJ: Lawrence Erlbaum Associates.

Widdowson, H. G. (1981). English for specific purposes: Criteria for course design. In Selinker, L., Tarone, E., & Hanzeli, V. (eds.), English for academic and technical purposes: Studies in honor of Louise Trimble. Rowley, MA: Newbury House Publishers.

Zareva, A. (2005). What is new in the new TOEFL-iBT 2006 test format? *Electronic Journal of Foreign Language Teaching, 2*(2), pp. 45-57.

Appendix A: Textbooks used in participating classrooms

Textbook Reference	Related Tasks and Page Numbers (if applicable)	Reference Information
Text A	Task 8 (pgs. 343-344, 347-348).	Edmonds, P, McKinnon, N. (2006). *Developing Skills for the TOEFL iBT: Intermediate.* Woodland Hills, CA: Compass.
Text B	Task 2 (pgs. 148-149)	Fellag, L.R. (2005). *Northstar: Building Skills for the TOEFL iBT Advanced.* White Plains, NY: Pearson Longman.
Text C	Task 7 (p. 275) Task 10 (p. 363)	Gallagher, N. (2007). *Delta's Key to the Next Generation TOEFL Test: Advanced skill practice book.* Surrey, UK: Delta Publishing Company.
Text D	Task 3 (pgs. 31-34) Task 4 (pgs. 217-218) Task 5 (pgs. 682-683)	Macgillivray, M., Yancey, P., & Malarcher, C. (2006). *Mastering Skills for the TOEFL iBT: Advanced.* Woodland Hills, CA: Compass.
Text E	Task 9 (pgs. 129-136)	Phillips, D. (2007). *Longman preparation course for the TOEFL test: iBT listening.* NY: Pearson Longman.
Text F	Task 1 (p. 26)	Wyatt, R. (2007). *Check your vocabulary for TOEFL.* London: A&C Black.
Internet A	Task 6	Davies, H. (2004). ESL conversation questions: Youth & old age. The Internet TESL Journal. Retrieved from http://iteslj.org/questions/age.html.