

A Comparative Washback Study of IELTS and TOEFL iBT on Teaching and Learning Activities in Preparation Courses in the Iranian Context

Shiva Seyed Erfani¹

¹Roudehen Branch, Islamic Azad University, Roudehen, Iran

Correspondence: Shiva Seyed Erfani, Assistant Professor, Roudehen Branch, Islamic Azad University, Roudehen, Iran. E-mail: Erfani@riau.ac.ir

Received: May 3, 2012 Accepted: May 22, 2012 Online Published: July 9, 2012

doi:10.5539/elt.v5n8p185 URL: <http://dx.doi.org/10.5539/elt.v5n8p185>

Abstract

One consequence of test use in the English-language teaching community is the negative washback of tests on teaching and learning. Test preparation courses are often seen as part of the more general issue of washback. IELTS and TOEFL iBT tests, focusing on communicative competence, are anticipated to have positive washback effect on how English is taught and learned. This study was a triangulation research to compare the washback of IELTS and TOEFL iBT on teaching and learning activities in test preparation courses. To this end, questionnaires survey with 40 IELTS and iBT teachers, 220 IELTS and TOEFL iBT students, and also ten class observations and ten teachers' interviews were called upon and the related data were analyzed quantitatively and qualitatively. The results showed evidence of both negative and positive washback in the relevant courses. Generally, in both courses the class time was more characterized by the negative influence of the test although IELTS classes seemed to promote more positive washback than TOEFL iBT did on class activities. While the TOEFL iBT preparation courses appeared to address a wider range of activities for academic study, IELTS preparation courses provided more opportunities for learners' interactions and communications in line with the test activities.

Keywords: IELTS, TOEFL iBT, triangulation, washback

1. Introduction

It is generally believed that tests have a powerful influence on language learners who prepare themselves to take the tests, and on the teachers who help learners to prepare. Alderson and Wall (1993) claim the existence of washback and believe that "... tests are held to be powerful determiners of what happens in the classrooms" (p. 41). Buck (1988) defines washback as "a natural tendency for both students and teachers to tailor the classroom activities to the demands of the test, especially when the test is particularly important for test takers" (p. 17). For Messick (1996) a poor test may be associated with positive effects and a good test with negative effects because of other things that are done or not done in the preparation classroom.

In 2006, the introduction of TOEFL iBT emphasized the need for testing experts, researchers, and teachers to realize that students should not only understand English but also communicate effectively to succeed in an academic environment in which English is the medium of instruction. As reported by Educational Testing Service (2008) "the innovations in the TOEFL iBT test and the emphasis on communicative competence are hoped to have positive impact on the way English is taught in preparation courses in which activities should more closely resemble communicatively-oriented pedagogy.

On the other hand, IELTS claims to measure test takers' communicative competency in all four language skills, and to predict one's ability to use English in genuine contexts. Zhenhua (2008) believes that IELTS is a more comprehensive language proficiency test with a better washback effect than iBT.

Many candidates take either IELTS or iBT in Iran on a regular basis and some attend relevant preparation courses to obtain required scores for higher education, immigration, graduation, employment, promotion, etc. The present study was carried out to compare the ways IELTS and TOEFL iBT influence teaching and learning activities, and sought their impact on the way English is taught and learned in preparation courses in the Iranian context. A comparative element was included since it is a common feature of washback studies (e.g., Alderson & Hamp-Lyons, 1996; Cheng, 1999; Hayes & Read, 2004; Shohamy, Donitsa-Schmidt, & Ferman, 1996; Wall &

Alderson, 1993; Watanabe, 1996).

The study looked at the phenomenon by an in-depth investigation of washback of the two tests to see whether IELTS and iBT functioned in the preparation classes the way the designers anticipated. It sought for a rational argument based on some empirical evidence to explore whether the same format preparation took place in IELTS and iBT courses. It also aimed to address whether teaching and learning activities were diverted from mainstream, well-designed language proficiency classes built around communicative competence and language skills required for academic settings and genuine context, into unproductive test taking strategies and sample test exercises. To fulfill the purpose of the study, the following research questions were formulated:

1. Is there any significant difference between teaching activities in IELTS and iBT preparation courses?
2. Is there any significant difference between learning activities in IELTS and iBT preparation courses?

2. Method

2.1 Participants

The participants were comprised of 100 IELTS students, 120 TOEFL iBT students, 20 IELTS and 20 TOEFL iBT teachers who took the questionnaire survey. Five IELTS and five TOEFL iBT teachers participated in the teachers' interviews. The same teachers' classes were also observed.

2.2 Instrumentation

The instruments included the students' questionnaire for those attending in preparation courses, the teachers' questionnaire for those preparing students for the tests, class observation table based on the questionnaires' items, and the teachers' interview including a set of questions about their teaching activities under the influence of IELTS or TOEFL iBT tests.

The students' questionnaire constituted 27 closed items, provided information about learning activities in preparation courses. The teachers' questionnaire also consisted of 16 closed items collected data on their teaching activities (the number of items remained after pilot study). All the closed items were on five point Likert scale of frequency. The checker items, restated in slightly different forms to evaluate consistency in responses, were also included. The items were mostly taken from Hawkey (2006) and Green (2007) since there were compatible with the purpose of this study. At the end of the questionnaires the students and the teachers were asked to write their comments on the courses.

2.3 Procedure

At the outset of the study a pilot study for the questionnaires was conducted with 30 IELTS and TOEFL iBT students and 5 IELTS and TOEFL iBT teachers which led to the elimination and modification of some items. Then the participants of the study were asked to complete the questionnaires. In terms of questionnaire development, qualitative input ensured the content validity. Factor analysis and Cronbach's alpha were also employed to ensure the construct validity and reliability of the questionnaires. Each set of items of both questionnaires met the three assumptions of singularity, sampling adequacy, and multi-co linearity. The selected response options were numerically coded and entered into a computer data base for descriptive statistics and chi-square analysis through SPSS version 18.

The comments of the participants at the end of the questionnaires were analyzed twice for two different purposes; first, to identify the relevant categories and, second, to count the frequency of occurrences which belonged to each of the derived categories. Then, the most frequent categories were reported.

Then five IELTS classes of the Academic Training Reading, Academic Training Writing, General Training Writing, Speaking, and Listening, and also five TOEFL iBT classes of Reading, Speaking, Listening, Writing, and Vocabulary were observed, video and audio recorded with CCTV cameras.

The observation task was divided into three tasks, involving pre-observation talks, recording events, and post-observation discussion. Prior to the observations the teachers were asked about the number of the session and the skills they taught, as well as the teaching materials they used. During the observations, various classroom events were written down in detail on a note-taking sheet. Upon the completion of each observation session, the teacher was asked about his or her intentions behind a variety of classroom activities. The purpose was to gather information that the researcher used to interpret the observation data. The types and contents of questions asked varied greatly depending upon what had been observed. Each observation session lasted for about 100 minutes. Like any triangulation research, as a part of validation process, class observations made a useful check on whether the IELTS and the TOEFL iBT classes displayed features similar to those described in the teachers' interviews and questionnaire responses by the teachers and the students. The transcriptions of classroom events

were analyzed to verify their claims on class activities in their preparation courses.

The qualitative analysis of class observations involved reviewing the notes and filling in the necessary information through reading the notes, watching the videos, and listening to the voice recordings. The interpretation of the meaning of behaviors described was deferred until a later time, because it is believed that simultaneous recording and interpretation often interfere with objectivity (Best & Khan, 1989). The transcriptions of classroom discourse were analyzed to categorize the characteristic features of classroom events. Then the approximate duration of each test-related activity in the observation table was calculated as a percentage of total class time to examine the validity of the predictions.

Next, the teachers whose classes were observed were interviewed. Each interview was held in a one-shot session taking between 20-30 minutes. They were conducted in English and audio recorded. During the interview time, a friendly rapport was established and interviewer bias was tried to be avoided. All interviews were immediately transcribed after the interview sessions. Then the answers of both the IELTS and the iBT teachers to each item were categorized. The results were reported in extracted forms.

3. Results

After estimating reliability of the IELTS and TOEFL iBT teachers questionnaires items (Table 1), an analysis of chi-square based on teachers' responses to questionnaire items was run to investigate the first research question quantitatively (Table 2).

Table 1. Reliability Index for IELTS and iBT Teachers' questionnaires items

	Cronbach alpha	N of Items
IELTS	.86	16
TOEFL iBT	.84	16

Table 2. Chi-square Analysis of IELTS and TOEFL iBT Teachers' Responses regarding Their Teaching Activities

	Value	Df	Asymp. Sig. (2-sided)
Pearson chi-square	37.14 ^a	4	.000

a. 2 cells (20.0%) have expected count less than 5. The minimum expected count is 2.19.

Given that the observed value of chi-square (37.14) exceeded its critical value (9.49) at 4 degrees of freedom, it was concluded that there was a significant difference between the teaching activities in the IELTS and the TOEFL iBT preparation courses.

Table 3. Frequency and Percentage of IELTS and TOEFL iBT Teachers' Responses regarding Their Teaching Activities

		CHOICES					Total	
		NOT AT ALL	NOT REALLY	Undecided	QUITE A LOT	VERY MUCH		
TEST	IELTS	Count	5	13	39	96	132	285
		% within TEST	1.8%	4.6%	13.7%	33.7%	46.3%	100.0%
	IBT	Count	0	17	63	89	54	223
		% within TEST	.0%	7.6%	28.3%	39.9%	24.2%	100.0%
Total		Count	5	30	102	185	186	508
		% within TEST	1.0%	5.9%	20.1%	36.4%	36.6%	100.0%

As displayed in Table 3, the teaching activities of 80 percent (a figure always showing the combined value of quite a lot and very much, shown in bold) of the IELTS teachers were influenced by the test while for a lower

percentage of the TOEFL iBT teachers (64.1 percent, the combined value, shown in bold) the test influenced their teaching activities. In other words, the teaching activities of a higher percentage of the IELTS teachers were influenced by the test than those of the iBT teachers. As for nonparametric analyses, a comparison between the percentages of teachers' responses to each item of the questionnaire was necessary to validate the results of the teachers' interviews and the class observations.

Table 4. A Comparison between the Percentage of IELTS and TOEFL iBT Teachers' Responses regarding Their Teaching Activities

<i>Teaching Activities Items</i>	<i>IELTS Teachers</i>	<i>iBT Teachers</i>
1. I work on information about format of the test.	88.9%	85.7%
2. I spend time looking at previous papers.	76.5%	57.1%
3. I work on marking assignments and giving feedback to the Students in the form of test band scores.	72.2%	50%
4. I work on techniques of taking the test.	100%	92.9%
5. I work on developing my students' genuine language ability.	83.3%	57.1%
6. I arrange my teaching according to the textbooks.	77.8%	57.1%
7. I adapt my teaching to the contents of the test.	88.2%	78.6%
8. I arrange my teaching according to the language skills.	82%	57.1%
9. I arrange my teaching according to the language activities.	88.9%	64.3%
10. I work on activities similar to those on the test.	94.5%	64.3%
11. I ask my students to take practice tests in class.	83.4%	71.4%
12. I usually ignore subjects that do not contribute directly to the passing of the exam.	44.4%	50%
13. I emphasize the activities which requires students to participate.	83.3%	64.3%
14. I work on integration of the skills rather than isolated skills.	72.2%	64.3%
15. I work on communicative activities of the target language domain.	72.2%	50%
16. I work on communicative requirements of the test.	72.2%	61.5%

As displayed in Table 4, IELTS teachers' answers to the items related to teaching activities showed a higher percentage than those of the iBT teachers. A closer look at the percentages revealed a range of differences depending on each item.

Following estimating reliability index of IELTS and iBT students' questionnaires items (Table 5), an analysis of chi-square based on the students' responses was used to investigate any significant difference between the learning activities in IELTS and TOEFL iBT on learning activities (Table 6).

Table 5. Reliability Index for IELTS and iBT Students' Questionnaires items

	Cronbach alpha	N of Items
IELTS	.79	27
TOEFL iBT	.80	27

Table 6. The Chi-square Analysis of IELTS and TOEFL iBT Students' Responses regarding their Learning Activities

	Value	Df	Asymp. Sig. (2-sided)
Pearson Chi-square	204.34 ^a	4	.000

a. 0 cells (0%) have expected count less than 5. The minimum expected count is 186.18.

Given that the observed value of chi-square (204.34) exceeded its critical value (9.49) at 4 degrees of freedom, it was found that there was a significant difference between learning activities in IELTS and TOEFL iBT preparation courses.

Table 7. Frequency and Percentage of IELTS and TOEFL iBT Students' Response regarding Their Learning Activities

		TEST * CHOICES Cross tabulation						
		CHOICES						
		NOT AT ALL	NOT REALLY	Undecided	QUITE LOT	A	VERY MUCH	Total
TEST	IELTS	Count	94	242	746	1091	418	2591
		% within TEST	3.6%	9.3%	28.8%	42.1%	16.1%	100.0%
	iBT	Count	289	479	753	877	341	2739
		% within TEST	10.6%	17.5%	27.5%	32.0%	12.4%	100.0%
Total		Count	383	721	1499	1968	759	5330
		% within TEST	7.2%	13.5%	28.1%	36.9%	14.2%	100.0%

As shown in Table 7, for 58.2 percent of the IELTS students (combined value, shown in bold), the learning activities were influenced by the test. While for the lower percentage of the TOEFL iBT students (44.4 percent, combined value, shown in bold) the test influenced their learning activities in preparation courses. In other words, the learning activities for a higher percentage of the IELTS students were influenced by the test than those of the iBT students in the relevant preparation courses. A comparison between the percentages of students' responses also provided a basis to verify the analyses of teachers' interviews and class observation.

Table 8. A comparison between the Percentage of IELTS and TOEFL iBT Students' Responses regarding Their Learning Activities

<i>Learning Activities Items</i>	<i>IELTS Students</i>	<i>iBT Students</i>
1. We are working on content of the test.	64%	61.4%
2. We are working on format of the tests.	83.7%	57.8%
3. The syllabus of the class is highly influenced by the test content.	57.3%	75.7%
4. We are taking practice tests.	73.4%	63.2%
5. We are receiving feedback in the form of test band scores.	55.3%	27.5%
6. We are working on test taking strategies.	70.1%	48.2%
7. We are working on test method activities.	84.6%	57.6%
8. We are doing tasks that require problem-solving.	55.2%	27.7%
9. We are working on knowledge creation and critical thinking.	45.9%	21.9%
10. We spend much of the time on pair and group work.	29.2%	9.5%
11. We have opportunities to participate and ask our questions.	75%	80.4%
12. Teacher dominates talking time.	57.3%	54%
13. We have opportunities to create our own interactions.	53.6%	38.8%
14. The class is based on communication-oriented teaching.	74%	28.5%
15. The class is based on score-oriented teaching.	39.6%	27.7%
16. We are working on development of language knowledge.	52.1%	53.9%
17. We are working on the development of academic skills.	66.3%	66.6%
18. We are working on the sentence level skills.	44.3%	33.4%
19. We are working on memorized language.	37.1%	60%
20. We are developing communication strategies.	53.6%	18.4%
21. We are developing skills with a use beyond the course.	43.6%	16.2%
22. We are working on integrated skills.	55.7%	37.3%
23. We are working on interactive language practice.	51.6%	19.2%
24. We are being trained just for test requirements.	62.3%	68.5%
25. We are developing study skills.	62.2%	42.6%
26. We are developing language skills.	65.2%	46.9%
27. Generally, skills development is influenced by the test.	69.8%	72.1%

Table 8 suggests that the IELTS students' responses related to the learning activities generally showed a higher percentage than those of the iBT students. In some items the differences indicated a wider range than in the others.

3.1 The Results of the Teachers' Comments Analysis

The results of the teachers' comments on their preparation courses at the end of the questionnaires revealed that the IELTS teachers mostly believed that although the courses provided opportunities for the students to get the required scores in a short time, the courses did not help them to learn English. They assessed the courses highly test-oriented which emphasized memorization rather than communication. The iBT teachers held that learning and teaching were both highly influenced by the iBT test. They believed that the students needed shortcuts like test taking strategies to get better scores. They claimed that the learners did not care about communication skills and they were simply after a passing score in a short period of time. Therefore, the teachers were pressed by the time limit. They also stated that the courses did not improve the students' language proficiency.

3.2 The Results of the Students' Comments Analysis

Both the IELTS and the iBT students mostly believed that the courses did not help them to become proficient in English language and what they learned was a number of strategies used for test taking. However, they assessed these courses useful for taking the test. The iBT students believed that they could improve their vocabulary although they had to memorize the items out of context. The IELTS students noted that the skills should have been simultaneously taught. Both the IELTS and the iBT students pointed out that they needed more opportunities to interact.

3.3 The Results of the Teachers' Interviews Analysis

Following the transcriptions of the teachers' interviews the categories relevant to the washback for each question were extracted as discussed below:

Extract 1. As to the difference between teaching IELTS/iBT courses and other courses they taught, most IELTS teachers claimed that they were specifically preparing their students for the exam and helping them to get the required score. However, they noted that in other courses they had more time for the students' interaction and actually they taught them English.

The iBT teachers mostly claimed that they prepared the students for the exam, taught them skills, techniques and strategies to take the test and emphasized time management. They also noted that they used Persian as the medium of instruction. However, they pointed out that in other classes they taught the English language to their students using English as the medium of instruction.

Extract 2. Considering their idea about the test format, generally both the IELTS and the iBT teachers perceived iBT more valid through its tasks of integrated skills, while they believed separated skills in IELTS affected the authenticity of the test. One IELTS teacher assessed IELTS test acceptable and authentic. However the IELTS teachers complained that the content of the test was problematic since the IELTS candidates needed a vast knowledge on different subjects to take the test. They perceived Academic reading and writing modules more difficult and listening and speaking modules easier for candidates.

Some iBT teachers held that the iBT was challenging because of its integrated task designed specifically to prepare candidates for academic settings. For some others integrated skills made the iBT format difficult. Some noted that although speaking to a computer was not natural, it had the advantage of reducing the stress of a face to face interview. Generally both groups assessed speaking to a computer screen as unreal, unauthentic and totally annoying and they found the IELTS speaking section more humanistic and authentic.

Extract 3. Regarding the easiness and difficulty of the IELTS/ TOEFL iBT test, all IELTS teachers agreed that the reading part was difficult which could not be answered without resorting to the test taking strategies. They perceived the speaking part easy and that the students got the highest score on this part. They also believed that IELTS test format was easy to learn but producing acceptable answers to fit that format appeared to be difficult. However, they believed that IELTS motivated the students to learn English. They regarded the British or the Australian accent used in IELTS as a weak point. The time pressure imposed by the IELTS test caused the students stress.

The iBT Teachers mostly believed that the easiest thing about the iBT was the possibility to get the required score through using test taking strategies. Obvious format, predictable content, preferred American accent, and the opportunity for note-taking throughout the test were considered as easy things about the iBT. They regarded overexposing of the students to a large number of vocabulary items and getting high score in spite of their unreadiness for academic or genuine settings, as weak points of the iBT test.

Extract 4. As what they hoped their students would learn from the preparation courses, the IELTS teachers mostly wished their students to succeed in the test through learning the techniques. However, some teachers

hoped that their students would learn creativity, communication, and developing authenticity of their language in a real context.

For the iBT teachers, note-taking, topic familiarity, time management and test-taking strategies were the things they wished their students would learn. They believed that the iBT preparation course was not a good choice to improve the English language.

Extract 5. In relation to *the activities they preferred to have more and the activities they preferred to have less*, the IELTS teachers claimed they preferred to improve their students' general English, to make their students communicate in pairs and groups, to teach them learning strategies, and to improve their critical thinking. The teachers complained that teaching reading comprehension was frustrating.

The iBT teachers stated that they preferred to teach test taking strategies as well as note-taking. They also emphasized that they preferred to focus more on the fluency of students and to teach them vocabulary in context; practically, they were unable to do so since it was not required by the course.

Extract 6. Concerning *the most and the least useful parts of the preparation courses*, the IELTS teachers considered test taking strategies and time management as useful parts. They also considered the development of reading and writing skills through e.g., paraphrasing, brainstorming, and outlining, useful and said that de-contextualized vocabulary teaching was useless.

For the iBT teachers, note-taking and test-taking strategies, time management, test format and topic familiarity were the most useful parts while vocabulary memorization was the least useful.

Extract 7. In relation to *the influence of the test on their teaching materials*, all IELTS teachers were convinced that the test influenced their choice of content in preparation courses. They mostly used past papers as their teaching materials. The IELTS teachers did not regard the IELTS preparation materials as a useful resource for improving students' proficiency.

The iBT teachers stated that although they had a variety of materials to choose from their teaching materials were influenced by the test.

Extract 8. About *the influence of the test on the choice of the methodology*, all IELTS teachers claimed that their choice of methodology was influenced by the test to some extent since they had to cover the materials in a short period of time. They also believed that the test had a guiding but not a dictating role in the way they taught. The teachers claimed that they benefited from pair and group work practices as well as collaboration and interaction to get their students involved.

All the iBT teachers claimed that the choice of their methodology was influenced by the test. They believed that since preparation courses were exam-oriented, they had to teach the students test taking strategies to increase their chance of passing the exam. They said that time pressure confined them to teach very specific points.

Extract 9. Regarding *the degree of success of their IELTS/iBT preparation courses compared with other English classes they taught*, the IELTS teachers mostly considered their courses more successful, because they were able to get a feedback after a short period of time. They also found the courses stressful but motivating.

Almost all the iBT teachers regarded their preparation courses more successful since they were able to see the results. They assessed the courses purposeful as well.

Extract 10. As to *the frequency of teaching an item they reckoned as important but it did not appear in the test*, most of the IELTS and the iBT teachers responded that they avoided teaching it.

Extract 11. When the teachers were asked to add *any comments concerning the influence of the test*, the IELTS teachers said that IELTS benefited from an international credibility. It was much more popular than iBT both for immigration or educational purposes. IELTS was motivating and made people pay more attention to English. They believed that iBT was not realistic. Reading on the screen and speaking to computers were not real. The only advantage of iBT over IELTS was the iBT integrated skills. However, they mostly believed that if candidates got good scores in IELTS it didn't mean that their English was good.

The iBT teachers believed that TOEFL iBT was a better test comparing to the IELTS. However, they noted that getting a high score on iBT was not an indication of the language proficiency. To these teachers iBT needed a lot of preparation and it was good for academic contexts while IELTS was a better indicator of the communicative skills and that culture bound questions made iBT biased. They convinced that the iBT was practical, and it expanded their knowledge of vocabulary.

3.4 The Results of the Class Observations Analysis

The comparison of the approximate percentage of time spent on each of the test-related activities in both IELTS and TOEFL iBT preparation courses shown in Table 9 reflected substantial differences between the courses.

Table 9. Test-Related Activities as a Percentage of Total Class Time

<i>Behavior Observed</i>	<i>Approximate IELTS Classes</i>	<i>Approximate iBT Classes</i>
1. Giving instruction on content of the test	7	6
2. Giving instruction on format of the test	4	8
3. Working on assignment and giving feedback on the correct forms	7	7
4. Giving feedback on the test scores	4	3
5. Working on sample tests and test answers	15	11
6. Working on test taking strategies	13	11
7. Working on skills development through test activities	15	7
8. Working on integrated skills	1	7
9. Working on test activities/tasks interactively	4	0
10. Assigning students in pair and group for test activities	5	0
11. Asking questions and students participation	7	6

Table 9 reveals that the IELTS courses spent more time on the test related activities than the TOEFL iBT courses did except for the time spent on giving instruction about the format of the test (e.g., the features included in the instructions for every section of the test) and working on integrated skills (e.g., writing on a topic after listening to a text or reading it). Meanwhile, both courses spent equal amount of time working on assignment and giving feedback on the correct forms.

The differences between the IELTS and the iBT courses in the amount of time spent on giving instruction about the content (e.g., the types of passages and the topic used in each section), giving instruction about the test scores (e.g., the criteria used for scoring each section), and working on test taking strategies (e.g., how to narrow down possible answers) were worthy of note. Although the differences between the two courses were much more evident regarding working on sample tests and test answers and also working on skills development (e.g., outlining, note-taking, paraphrasing) through test related activities. The surprising differences between the two courses were the amount of time spent working on test task interactively and assigning students in pair and group for test activities in the sense that while there were substantial amount of these activities in IELTS courses they were completely absent in iBT courses.

In addition to the test-related activities included in Table 9, several other points were also observed. The most immediately observed difference was that the medium of instruction in IELTS courses was English and teachers encouraged students to speak English in the class while the medium of instruction in TOEFL iBT courses except in Speaking class was Persian.

The IELTS teachers provided many opportunities for the students to interact in pair and group and to do communicative activities in line with the test task. In TOEFL iBT preparation courses teachers dominated talking time, thus there was too little opportunity for the students to create their own interactions which left students inactive as well.

With regard to the source of classroom materials, published IELTS preparation texts by Cambridge University Press mostly including past papers were the predominant source of IELTS classes. In TOEFL iBT classes, they made use of a variety of text books such as TOEFL iBT Longman, Barron's, ETS Guidelines, Sample Tests, and supplementary provided by the language school. Therefore iBT courses provided a wider range of test related activities required by academic setting.

Both IELTS and TOEFL iBT teachers in the speaking classes rarely corrected some pronunciation, grammar and vocabulary errors while students were in the process of communicating. This may be acceptable to maximize communicative opportunity. In post observation talks it was found that the teachers believed the treatment of errors resulted in inaccurate oral communication even to the point of impeding meaning. This finding is in line with the results obtained by Hawkey (2006).

Teachers in both courses also spent a portion of time teaching language knowledge. However, iBT teachers went

for deductive teaching of grammatical rules and vocabulary using explicit explanations. Instances of memorized language were observed in both courses; IELTS students were required to memorize a few prefabricated chunks of language such as proverbs, idioms and vocabulary. TOEFL iBT students were extensively asked to memorize exhausting lists of de-contextualized vocabulary and grammatical rules.

There were a few practices of critical thinking and problem solving activities (e.g., asking for different interpretations of a text) parallel with test related activities in IELTS classes. Creativity was also evident in IELTS courses when the speaking teacher asked the students to imagine, create and then talk when each student was obliged to say something new in the task. However, there was no sign of such activities in iBT courses; in iBT speaking class a cliché was dominant and each student followed it to complete a sample test task while one of the students was controlling the time. In the post observation talks with the teacher, she suggested that the candidates were not going to be assessed at the same time, therefore learning a correct pattern was more useful than creativity to pass the exam.

The teachers in both courses repeatedly provided the students with feedback on how to manage their time during the test administration. This was more evident in iBT than in IELTS classes.

4. Discussion

It was not expected to find either totally positive or totally negative washback in any of the preparation courses, given that the reality of washback is bound to be more complicated.

Questionnaires survey, teachers' and students' comments, teachers' interviews and class observations showed that there were substantial areas of common practice between the IELTS and the TOEFL iBT preparation courses; IELTS and TOEFL iBT were perceived as high-stakes tests, motivating both candidates and teachers, although most of the time to the extent of demands for narrower test-based preparation courses. Teachers of IELTS and iBT courses shared the belief that the design of the tests dictated practices in preparation courses which focus on the range of relevant activities, skills and materials inspired by the content of the iBT and IELTS tests, therefore most of the class time in both courses spent working on sample test and past papers. This finding is supported by Hawkey (2006) in his IELTS impact study.

Both class types focused extensively on test format, test taking strategies, test content, getting high scores on the test and time management. These activities are characterized by negative washback as concluded by Watanabe's study (2004) on teacher factors mediating washback.

Teachers and learners of both courses regarded the test preparation courses in instrumental terms. The purpose of the courses was to ensure that students achieved the required grade. They were all concerned with strategies for passing the test rather than developing language proficiency as their comments suggested. This finding supports the generalized assertions of Berry and Lewkowicz (2000) and Lumley and Stoneman (2000) about the nature of language test preparation courses in which learners only focus on practicing and mastering item types for the tests not on learning the language. This distinction is also claimed by the teachers of both courses in their interviews regarding the differences between the preparation courses and their other proficiency classes. This contrasts with the finding of Watanabe's (2000) study in which the teachers claimed that they deliberately avoided referring to test taking techniques since they believed that actual English skills would lead to students passing the exam. Nevertheless high percentages of the IELTS and the TOEFL iBT teachers of this study claimed in their questionnaires and interviews that test taking strategies rather than language proficiency was what the students needed to get the required score on the tests. Interestingly, this idea is also supported by the fact that teachers of both courses claimed that they usually ignored subjects and activities that did not contribute directly to the passing of the exam.

The courses under study were organized around the practice of skills, particularly through test-related tasks in different ways. There were some opportunities for communication, exchange of information and interaction, and assigning students in pairs and groups to perform test-related tasks in the IELTS preparation courses confirmed by both questionnaires survey and class observations. These activities were characterized by positive washback as described by Watanabe (2004). The development of the skills required by the test through skills building activities is supported by the findings of Green (2007) in his study on preparation for academic writing in higher education. Teaching materials used in IELTS courses basically included past papers which narrowed the range of activities performed in these classes. However, the use of a variety of textbooks in iBT classes provided a wider range of test-related activities required for an academic setting of English speaking countries.

Questionnaires survey, teachers' interviews and observation of classes all suggested that IELTS and TOEFL iBT affected both what and how teachers teach. This finding supports the result of research by Alderson and Hamp-Lyons (1996) that the TOEFL affected both what and how teachers teach but contradicts the conclusion

drawn by Alderson and Wall (1993), Watanabe (1996) and Green (2003), whose observations of IELTS preparation and EAP classrooms indicated that course content was very clearly influenced by the test, but any influence on teaching and learning method was less obvious.

Although the study found that the teachers' and the students' claims on some items of the questionnaires were contradictory, the class observations analysis revealed that the IELTS and iBT teachers responses in their interviews accurately reflected their teaching.

5. Conclusion and Pedagogical Implications

This study concluded that the teachers and the students in both preparation courses generally focused on practice of the test tasks, rather than the development of academic language proficiency in a broader sense. This conclusion is in line with the one by Hayes and Read (2004).

Generally, this study found evidence of both negative and positive washback for the IELTS and the TOEFL iBT test. This finding is supported by Por and Ki (1995) that both positive and negative washback tend to be associated with high-stakes testing. However, the amount of each type of washback between the two tests varied greatly in different aspects.

The positive washback found in both the IELTS and the iBT courses was the promotion and development of skills through test related activities, asking students for participation and interaction while doing test tasks; however these activities were more evident in the IELTS courses. Among other positive washback effects seen in this study was working on integrated skills in iBT courses. On the other hand, working on communicative activities and assigning students to pairs and groups to perform the test activities were exclusively found in IELTS courses.

The negative washback effect found in this study both in the IELTS and the iBT courses were narrowing of the scope and content of teaching and learning. This kind of negative washback has been referred to in the language testing literature as narrowing of the curriculum (Madaus, 1988). In this study the negative washback was also reflected in students' memorizations of vocabulary and other materials rather than acquisition and development of skills overwhelmingly seen in the iBT courses which could lead to the increased test scores without a parallel rise in ability in the construct tested, referred to in literature as test score pollution (Haladyna, Nolan, & Hass, 1991). Another negative case of washback evident in the teachers' interviews and class observations was little attention paid to the materials not included in the tests. This evidence is also considered as negative washback in the study of Wall and Alderson (1993).

A number of implications emerged from this study at the Pedagogical level. The study highlighted the significance of appropriate preparation courses. Test administrators, curriculum coordinators, and language institutions should be aware whether the impact desired by the test designers is the way it should be in preparation courses. They should provide teachers with the guidance necessary to present the materials in more effective ways to reduce negative washback and to familiarize them with not only the test format but also the constructs underlying the test design. Teachers should be provided with the knowledge to teach communicatively to promote positive washback in preparation courses.

Language testing researchers should be much more concerned about test preparation practices, a problem not related to language testing but to program administration, to teaching, to textbook, and to the educational-commercial interface, as suggested by Hamp-Lyons (1998). Although the research on preparation courses may not solve the whole issue of test washback on teaching and learning practices, it can help through a better understanding of the phenomenon.

Curriculum developers and teachers should be aware of the general assertion that narrowing of the curriculum in response to test demands contributes to distortion in the interpretability of test results (Green, 2007). As Green suggested, by focusing on features of the test rather than on the focal construct, it may be possible to improve scores without improving target abilities.

Students must have sufficient understanding of the design of the test and its implications for their learning. The students may assume that reaching the level required for the test indicates their readiness for academic or vocational settings; however, they should be ensured with the opportunities to go beyond the demands of the test.

Further investigation on the attitudes of teachers and students, the role of textbooks and the effectiveness of the courses on test performance and the relationship between the candidates' scores on IELTS and TOEFL iBT and their future academic success must be carried out. As the number of studies of high-stakes proficiency tests like the IELTS and the TOEFL iBT preparation courses increases, we will gain a better understanding of the washback of the test in the way English is taught and learned in different classrooms and, more generally, its impact in a particular social context.

References

- Alderson, J. C., & Hamp-Lyons, L. (1996). TOEFL preparation courses: A study of washback. *Language Testing*, 13, 280-297. <http://dx.doi.org/10.1177/026553229601300304>
- Alderson, L. C., & Wall, D. (1993). Does washback exist? *Applied Linguistics*, 14(2), 115-129. <http://dx.doi.org/10.1093/applin/14.2.115>
- Berry, V., & Lewkowicz, J. (2000). Exit-test: Is there an alternative? *Hong Kong Journal of Applied Linguistics*, 5(1), 19-49.
- Best, J. W., & Kahn, J. V. (1989). *Research in education*. Prentice-Hall, Inc., Englewood Cliffs, NJ., USA.
- Buck, G. (1988). Testing listening comprehension in Japanese university entrance examinations. *JALT Journal*, 10, 12-42.
- Cheng, L. (1999). Changing assessment: Washback on teacher perspectives and actions. *Teaching and Teacher Education*, 15, 253-271. [http://dx.doi.org/10.1016/S0742-051X\(98\)00046-8](http://dx.doi.org/10.1016/S0742-051X(98)00046-8)
- Educational Testing Service. (2008). *Validity evidence supporting the interpretation and use of TOEFL iBT scores*. Princeton, NJ: Author.
- Green, A. (2003). Test impact and English for academic purposes: A comparative study in backwash between IELTS preparation and university pre-sessional courses. Unpublished PhD thesis, Centre for Research in Testing, Evaluation and Curriculum in ELT, University of Surrey, Roehampton.
- Green, A. (2007). *IELTS washback in context: Preparation for academic writing in higher education*. Cambridge University Press.
- Haladyna, T. M., Nolen, S. B., & Haas, N. S. (1991). Raising standardized achievement test scores and the origins of test score pollution. *Educational Research*, 20(5), 2-7.
- Hamp-Lyons, L. (1998). Ethical test preparation: The case of TOEFL. *TESOL Quarterly*, 33(2), 329-337.
- Hayes, B., & Read, J. (2004). IELTS test preparation in New Zealand: Preparing students for the IELTS academic module. In L. Cheng, Y. Watanabe, & A. Curtis (Eds.), *Washback in language testing: Research contexts and methods* (pp. 19-36). Mahwah, NJ: Lawrence Erlbaum Associates.
- Hawkey, R. (2006). *Impact theory and practice: Studies of the IELTS test and Progetto Lingue2000*. Cambridge University Press.
- Lumley, T., & Stoneman, B. (2000). Conflicting perspectives on the role of test preparation in relation to learning? *Hong Kong Journal of Applied Linguistics*, 5(1), 50-80.
- Madaus, G. F. (1988). The influence of testing on curriculum. In L. N. Tanner (Ed.), *Critical issues in curriculum*. Eighty-seventh yearbook of the University of Chicago Press.
- Messick, S. (1996). Validity and washback in language testing. *Language Testing*, 13(3), 241-256.
- Por, L. H., & Ki, C. S. (1995). Methodology washback: An insider's view. In D. Nunan, R. Berry, & V. Berry (Eds.), *Bringing about change in language education* (pp. 217-235). Hong Kong: University of Hong-Kong, Department of Curriculum Studies.
- 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>
- Wall, D., & Alderson, L.C. (1993). Examining washback: The Sri Lankan impact study.
- Wall, D., & 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 Research Report. TOEFL iBT-05. 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, 318-333. <http://dx.doi.org/10.1177/026553229601300306>
- Watanabe, Y. (2000). Washback effects of the English section of Japanese university entrance examinations on instruction in pre-College level EFL. *Language Testing Update*, 27, 42-27.
- Watanabe, Y. (2004). Methodology in washback studies. In L. Cheng, Y. Watanabe, & A. Curtis (Eds.), *Washback in language testing: Research contexts and methods* (pp. 19-36). Mahwah, NJ: Lawrence Erlbaum Associates.
- Zhenhua, Y. (2008). *An analysis of the increasing popularity of IELTS (International English Language Testing System) in China*. Australian Studies Centre, Renmin University of China.