Washback Of English Proficiency Test B1 In Classroom Activities At Hanoi Law University

Đinh Thi Phuong Hoa
Hanoi Law University

Abstract

A study in Vietnam concerning the effects of The Vietnam Six-levels of Foreign Language Proficiency Framework, specially English Proficiency Tests for graduates, on classroom teaching and learning activities are reported. The study explores the phenomenon of washback or backwash, the influences of testing on 9 teachers and 715 non-English major students. It is cited as the only known research investigating washback in language education through classroom observation. The study was conducted at Hanoi Law University, and combined classroom observations with data from interview, questionaire responses and document analysis to determine whether washback exist, to what degree it operates, and whether it is a positive or negative force in this educational context. The insights from the findings indicate that washback of English Proficiency Tests for graduates occurred in both positive and negative forms, to some degree, in teaching and learing content, methods and styles. Evidence of washback, both positive and negative, on the way teachers design tests was also found. This should help Vietnamese educators to prepare favourable conditions for enhancing the benificial washback of EPT. The findings have contributed to the knowledge of a nature of washback and consequently opened a new understanding to recognize the dissimilar levels of washback. further research is recommended.

Keywords: *washback, English Proficiency Tests, classrooms activities*

I.   Introduction

Today, English has become a global language that offers the chances to integrate into all the professions. Khamkhien (2010, p. 757) stated that, "the importance of English has flashed an increasing concentration in the development of English language teaching in numerous countries" [1]. In Vietnam, English has been instructed nationwide as a compulsory subject at both lower, upper secondary level and tertiary level; and as an elective subject at primary level from 1980s to present (Nguyen, 1997, p.5) [2]. Notwithstanding its impact, English language teaching and learning for non-major learners in Vietnam are contradictory to all expectations as the language skills of listening, speaking, reading and writing have not been appreciated. Furthermore, most of the teachers have still taught English with traditional techniques, as teacher-centered or the grammar-translation method for many years. Conversely, for fulfilling the needs of a modern society in the globalization epoch, Vietnamese Prime Minister issued Decision N° 1400/QD-Ttg of September 30, 2008, approving the scheme "Foreign Language Teaching and Learning in the national education system during 2008 – 2020" and now this scheme is extended to 2025 (National Foreign Languages Project for short) [3]. The scheme aims at implementing an educational innovation and evaluation of foreign language teaching and learning at all levels in the national education system. Accordingly, Minister of Education and Training issued the Circular N⁰ 01/2014/TT-BGDĐT of January 24, 2014, approving *The Vietnam Six-levels of Foreign Language Proficiency Framework* (henceforth VNFLPF). This framework consists of six levels that are compatible with the Common European Framework of Reference for Languages [4] (CEFR for short) and other common international language proficiency levels and used as reference when writing curriculums and teaching plans. According to National Foreign Languages Project, with undergraduate institutions that are not specialized in foreign languages, the new language-training program must require a language proficiency of VNFLPF level 3 upon graduation. Based on this framework, English Proficiency Test from level 2 to level 5 (henceforth EPT.2 and EPT.3-5) is conducted and issued. Among these, EPT.2 is compatible with A2 of CEFR and EPT.3-5 is compatible with B1, C1, and C2 of CEFR. It thus became a very high-stakes test with serious consequences for non-English major students.

Hanoi Law University (henceforth HLU) is Undergraduate institution that is not specialized in foreign languages; the new language-training program must require a language proficiency of VNFLPF level 3 upon graduation. However, because of limitation of training time and English in mixed-big sized classes (from 55 to over 65 students) in a large room where is 105 square meters in area, no microphone, and thus, some students

could not listen to all lessons clearly. 715 students were from 18 to 22 years of age. They were from different Northern areas of Vietnam. Although, they had 3 years of learning English at high schools, their English proficiency was at beginner level (A0), therefore, the Rector of HLU decided to apply for English proficiency of VNFLPF level 3 (B1) upon graduation. As a result, EPT.3 (B1) of VNFLPF is a compulsoty requirement for HLU graduation from 2018 .

On the basic of the background of the Vietnam educational innovation context, particularly the context at HLU, the study attemped to address these issues:

1) Whether English Proficiency Test will positively influence the English language teaching process at Hanoi Law University, Vietnam.

2) Whether the changes in the teaching process will beneficially affect teaching strategies, which will lead to changes in learning style at Hanoi Law University, Vietnam.

II. Literature Review

The definition of washback in this study

The term "washback" is predominant in language teaching and testing literature as well as general education. However, the term "washback" has been defined and interchangeably by many researchers and organizations worldwide.

In applied linguistics, the term "washback" or backwash is defined as the influence or impact of tests on curriculum/ syllabus design, language teaching and language testing [5]. Accordingly, tests can influence teachers and learners, and thus influence teaching and learning activities. The influences may be either positive or negative, depending on various facets not yet defined. Nevertherless, whether a seperate and distinguishable phenomenon of washback exists is still open to debate; and there appear to be very few emprical studies directly investigating this phenomenon [6].

In the educational evaluation literature, washback is considered the influences of testing on teaching and learning practices. Therefore, tests can drive teaching and learning that is also mentioned as measurement-driven instruction [7]. Fitz-Gibbon (1996) defined impact as any effect of the service [or of an event or initiative] on an individual or group [8]. This definition accepts that the impact can be positive or negative and may be intended or accidental. When holding this definition, measuring impact is about identifying and evaluating change [9].

Messick (1989) expanded the concept of consequential validity, changing the previous notions about score interpretation and test use. The concept of washback in test validity research is primarily associated with Messick's concept of consequential validity. Therefore, washback is defined as an "instance of the consequential aspect of construct validity and a focal point of validity research" [10], which covers components of test use, the impact of testing on test-takers and educators, the interpretation of results by decision-makers, and any possible misuses, abuses, and unintentional effects of tests. The influences of tests on teachers, students, institutions, and society are accordingly considered one type of validity evidence. Many other researchers have also emphasized the meaning of justifying test use and exploring its consequences ([11]; [12]). Therefore, washback also plays a key role in the process of educational innovation and assessment in language teaching and learning [13].

In short, for the purpose of this paper, the term "washback/backwash" is understood to be the influences that tests have on teachers and students in terms of the methods/activities they use in their classrooms to teach/study English as Foreign Language.

The Vietnam Six-levels of Foreign Language Proficiency Framework

The CEFR provides a detailed description of learner level by skills, in a language-neutral format. Therefore, the CEFR is used for many dissimilar practical purposes because its influence goes beyond merely describing language proficiency of learners, they are: teacher training programs, developing syllabuses, creating tests/exams, marking exams, evaluating language learning needs, designing courses, developing learning materials and describing language policies continuous/self-assessment.

Accordingly, VNFLPF is designed based on CEFR in the Vietnam educational context. This framework consists of six levels and its Can-do descriptors that are compatible with CEFR and other common international language proficiency levels. Therefore, VNFLPF is used as reference when writing curriculums teaching plans, assessement and designing test.

VNFLPF describes foreign language proficiency at three broad bands with six main levels: level 1 and level 2, level 3 and level 4, level 5 and level 6. The scale starts at level 1 and finishes at level 6 that is compatial with CEFR from A1 to C2 as the following:

Table 1: The 6 levels of the VNFLPF

VNFLPF (Level)		General Descriptions
A – Basic user	Level 1 (A1 – Breakthrough)	Can communicate in basic English with help from the listener
	Level 2 (A2- Way-stage)	Can communicate in English within a limited range of contexts
B – Independent user	Level 3 (B1 – Threshold)	Can communicate essential points
	Level 4 (B2 – Vantage)	Can use English effectively, with some fluency, in a range of contexts
C – Proficient user	Level 5 (C1 – Effective Operational Proficiency)	Can use English fluently and flexibly in a wide range of contexts.
	Level 6 (C2 -Mastery or Highly proficient)	Can use English, very fluently, precisely and sensitively, in most contexts

For the purpose of this paper, the usage of VNFLPF helps to define clearly certain requirements for competency, capacity in listening, speaking, reading and writing, and thus English level 2 (A2) of VNFLPF learner is actived in the performance of the four main language activities, including listening, speaking (spoken interaction), reading, writing (written production) in the public, the personal, the educational and the occupational domains with some types of text and questions.

Some washback studies

Studies on washback reveal varied and sometimes different findings. The following section discusses the washback influences on teaching or learning activities in classroom.

The field of washback has been investigated by many researchers around the world. Among these, the washback model of Alderson and Wall (1993) is considered a classic and landmark study. Alderson and Wall (1993) used obsevation method to carry out their Srilanka study on investigating the washback existing of English teaching and learning activities in classroom. Alderson and Wall (1993, p. 120-121) developed the fifteen hypotheses (WHs for short) that combined different possible aspects of washback, including the effect on what to teach/learn, how to teach/learn, the rate and sequence of teaching/learning, the degree and depth of teaching/learning and the attitudes to content, method, etc. of teaching/learning [14]. Alderson and Hamp-Lyons's model (1996, p. 296) used interviews and one-week-classroom observations of teachers to review and correct WHs of Alderson and Wall (1993) that "tests will have different amounts and types of washback on some teachers and some learners than other teachers and learners" [15].

The studies of Cheng (1999 and 2004) focused on old and new HongKong Certificate Examination in English (HKCEE) ([16]; [17]). Cheng (1999) used classroom observation that combined her data of baseline study and Part A of Communicative Orientation of Language Teaching [18] and interview methods to compare "teachers' perceptions toward both old and new HKCEE". Cheng (2004) based on a combined research framework that employed multiple approaches to explore both the macro level (including the main parties within the HongKong educational context) and the micro level in schools (concerning different aspects of English teaching and learning) to recognize the washback phenomena by using English questionaires that consisted of three parts. Part 1 discovered the general information of teacher. Part 2 with 5-point Likert scale of agreement discovered teacher's perceptions and 5-point Likert scale of frequency of Part 3 discovered teacher's reactions to the new HKCEE through their classroom teaching and learning activities.

Regarding the washback of CEFR, Pan and Newfields (2012) worked on discovering how English proficiency graduation requirements have impacted 17 tertiary educational institutions in Taiwan by using extensive questionnaire and interview data [19]. Among them, the survey contained two types of questions: multiple-choice questions with categorical responses and 5-point Likert scale questions with pseudo-ordinal responses. Since 2003, Taiwan's Ministry of Education (TME) established a list of recommended tests to set English thresholds for graduates to generate a level of English proficiency, which were modified according to the CEFR B1 or A2 levels. They included two local tests: the General English Proficiency Test (GEPT) and the College Student English Proficiency Test (CSEPT). The GEPT was a 5-level, four-skill general English proficiency examination commissioned by TME in 1999. The CSEPT was 2-level, listening-reading-grammar test for university-level students in Taiwan. Accordingly, Pan and Newfields (2012) conducted their study after the inception of English certification graduation requirements in Taiwan; so a comparison of the baseline and a follow-up study to determine the consequences brought about by the tests was not viable. Therefore, a comparison of the differences between the schools with graduation requirements and those without graduation requirements will be used to reveal test effects.

In short, this part focuses on some washback studies published between 1993 and 2012. The first part also reviews how these studies have investigated washback. All studies cited here explore different aspects of washback and use various instruments. Alderson and Wall investigated evident of both beneficial and harmful washback on the content of teaching and on ways of assessing, but not on teaching methodology. Alderson and Hamp-Lyons ascertained the influence of the TOEFL on class teaching and TOELF affected both what and how

teachers teach, but the effects differed from teacher to teacher. However, the study of Alderson and Hamp-Lyons had three significant limitations. Firstly, they did not include questionaires. Secondly, they choiced participants and lastly, they dealt with washback primarily from perspectives of teachers, hardly addressing students' opinions. Cheng contributed to the few washback studies by using both quantitative and qualitative methods. Cheng's study was useful because her study attemped to evaluate the effects of the new examination, however, a longitudinal research with a longer timeframe than the one used by Cheng might shed better light on the influences of the new HKCEE. Pan and Newfields aimed to discover the test effects brought about by graduation requirements in the Taiwanese tertiary educational framework from the perspectives of students. Comparatively little research of Pan and Newfields was conducted regarding the effects of tests on the learning processes, in contrast to the significant number of studies on the effects of tests on teaching. Their study confirmed the argument of other washback studies that standardized tests were not a panacea that always succeeded in changing students' study habits. Therefore, their study also made it clear that the test requirements did not lead to a notable amount of "*studying for the test*" a phenomenon often reported in examination-oriented societies. However, the study of Pan and Newfields had three noteworthy limitations. Firstly, the study of Pan and Newfields was conducted at a period when many institutions in Taiwan were eager to adopt the government's EFL graduation exam policy and thus, washback appeared to be ineffective. Secondly, their study has relied on self-reported student data and thus, such information was easily prone to expectancy bias. Subsequent investigation should include more classroom observational data and seek to corroborate student data with other data sources from teachers and school administrators. This should allow their study to get a more accurate and dynamic picture of how washback patterns are perceived by different test stakeholders. Lastly, one goal of introducing graduation requirements was to improve the ability of graduates to communicate effectively in English in the office that would be very difficult for the researchers to measure, further research should pay more attention to this aspect of washback.

All of the reviewed studies have been conducted in primary and secondary schools or tertiary educational institutions in Srilanka, HongKong, Taiwan. The methods were used involved either written questionaires or interview/observations. They found evidence of washback influences on teachers' behaviours or learning. Accordingly, there has no previous research into washback effects arising from EPT.2 or standardised tests in Vietnam Hanoi Law University. Thus, further research into this area is still needed.

Drawing on some washback models and some empirical studies on language teaching or learning activities in classroom of Alderson and Hamp-Lyon (1996), Cheng (1999 and 2004) and Pan and Newfields (2012), this study will be designed to investigate "*Washback of English Proficiency Test in Classroom Activities at HLU*". The study concentrated on an exploration of the effects of washback on teachers/teaching process may offer insights about how VNFLPF and EPT.2 influence language teaching or learning activities in classroom at the educational innovation of HLU context. Furthermore, evidences from various sources of this study also helped to consider how the teachers and students benefit from the innovation.

III. Methodology And Data

This work was conducted between January 2014 and November 2020, aiming to capture the changes when VNFLPF was introduced into teaching in 2014 until the first cohort of HLU's students took the EPT.2 graduation examination in 2018.

For ensuring the validity and reliability of the questionnaire items, qualitative input and piloting procedures were carried out that lead to ensure the content validity and thus its consequential validity [20]. This study collected data of three kinds: (1) documents analysis, (2) focus group interview, (3) questionaires and (4) classroom observations. However, the data from (1), (2) and (3) are considered backdrop to the discussion (4) because of the extent of the data and space limitation.

Subjects of the study

The subjects of the project were Rector of HLU, Head of Training Department, 14 teachers of English at HLU (02 Vice Directors of Foreign Language Central and 9/14 teachers of English) and 715 non-English major students of HLU.

Conducting the document analysis

The researcher collected all institutional policy documents on innovating methods of assessment, syllabus, and supplementary materials according to VNFLPF and EPT.2 for getting the data because such artifacts of everyday experience can provide information about what has been encouraged or discouraged; about what has happened or will happen ... etc. [21]. Therefore, such documents are particular useful for educational research.

Conducting the questionaire

The survey of this study was carried out within from December 25, 2017 to January 12, 2018. Simple random sampling was employed in this study. For comparing the correct responses given by each group, Teacher Questionnaire and Student Questionnaire consisted of four parts and the same contents that were modified and adapted to Cheng (2004). All items of Questionnaires were designed according to the results of VNFLPF and EPT.2 analysis. The same contents of Teacher Questionnaire and Student Questionnaire was designed to check who remember or who tell the truth and thus, determine what happens in classrooms activities and how washback operates if it occurs.

Due to the length of this study, Teacher Questionnaire and Student Questionnaire were described shortly as the following:

Teacher Questionnaire and Student Questionnaire

Numerical order	Concepts	Variables	Scales
A: Personal details			
Part 1	Engling proficiency, ages	2/4	Nominal Scale
B. EFL teaching and learning activities:			
Part 2	Contents and communicative method of teaching EFL (including listening, speaking, reading, writing skills)	372	5-point Likert scale of frequency
While-lesson activities	Topics	52	
	Texts	72	
	Question types	32	
	Activities	28	
After school (Homework)	Topics	52	
	Texts	72	
	Question types	32	
	Activities	28	
Post-lesson activities	Correct and Comment	4	
Part 3	Materials	13	Nominal Scale
Part 4	Assessment (including listening, speaking, reading, writing tests)	10	Nominal Scale

Conducting the observations and instruments

After receiving the permission of all participants, 10 classes (English level A2) of ten teachers were chosen for observing. The researcher conducted the observations to obsever what happens in the English classroom and thus, determine what and how teacher teach or what and how students learn.

The observation happened from January to March 2018. Classes are scheduled one day per week with substantial uninterrupted work periods and the teaching session lasted approximately 200 minutes (4 periods) per day every morning or afternoon. There are 55 periods of English level A2 from December 25, 2017 to March 23, 2018. Therefore, the observation process was divided into 2 rounds, they were Round 1 and Round 2. Round 1 was took palace that far from the semeter examination ans Round 2 was observed before the semeter examination to explore the differences of influences of VNFLPF and EPT.2 between two Rounds.

50 minutes of every observation was the length of each lesson period and the teaching session lasted approximately 200 minutes (4 periods) per day every morning or afternoon . The observation process was divided into 2 rounds, they were Round 1 and Round 2 as the following:

Table 3.4 Observation Timeline

Duration: Spring semester, 2018		
Round 1	**The length of classroom observation period**	**Time**
10 English lessons	50 minutes for each observation of one English lesson	from January 5 to March 23, 2018
Round 2	**The length of classroom observation periods**	**Time**
30 English lessons	150 minutes for each observation of 3 English lessons	from March 26 to March 30, 2018

For getting the exact information, thick descriptions and the responds of teachers and students in a natural manner, teachers and students were explained about the observation. The observation scheme was designed and adapted according to Cheng (1999) that combined the data of mentioned questionaires and Part A of COLT. These descriptive data would be assessed according to their common outcomes. Therefore, the researcher also discovered how VNFLPF and EPT.2 influences teachers and students.

Conducting the interviews

After observations, the focus-group interviews were held because the participants had a few experiences of teaching and learning English by that time. The open-ended questions were designed to attain the best feasible quality of responses from the members because the open-ended questions were used to add the depth of the data via participants' individual experiences [21]. Moreover, these open-ended questions were applied in both individual and focus group interviews. This combination helped to focus on getting the specific information that would be comparable across the group of participants.

The researcher carried out at least 05 minutes of some focus group interviews for triangulation after observation. Among these, the researcher took note the attitudes of the teachers and students and the discussion between the participants when taking tasks were allocated in order to discover what teachers used and taught, and how students responded. After receiving the permission of some participants, some formal focus group interviews were audiotaped and transcribed in short, and thus translated precisely.

The analysis procedures

The analysis of involved a calculation of the amount of time/times was applied to the observation data and Part 1, 3 and 4 of Teacher Questionnaire and Student Questionnaire by using Excel and IBM Statistical Product and Services Solutions software. The survey explored the differences between findings of two groups (Teacher and Student). The differences were tested for determining statistical significance by using the Lavene's test and the independent sample T-test. The Levene's test for Equality of Variances was used to clarify the equal distribution in each subgroup. The Independent Samples Test compares the mean scores of two groups on all given variables. A probability of less than .05 was taken as statistically significant for the survey ($p < 0.05$). If it is not significant, the value is greater than .05 ($p > 0.05$), the two variances are not significantly different; that is, the two variances are approximately equal. If the Levene's test is not significant, the second assumption should be met. The possibility of error could increases with the number of T-tests being carried out. Accordingly, a method triangulation with a complementary multiple-method design were used in this study to ensure against errors arising from the data collection and analysis. The present study was designed after the beginning of English graduation requirements at HLU; so a comparison of the baseline and a follow-up study to define the consequences brought about by VNFLPF and EPT.2 were not viable. Therefore, a comparison of the differences between teachers and students will be used to reveal test effects in the classroom activities as the following findings and discussion.

IV. Findings And Discussion

Results of Document Analysis

As stated in the methodology, document analysis involved institutional policies on curriculum, the official course documents, methods of assessment and supplementary materials used by teachers. Relevant details of the analyses are given below.

Curriculum and methods of assessment

(a) Teaching contents and methods of assessment have been changed. *Table 4.1.1* illustrates the changes in teaching contents and methods of assessment

Table 4.1.1 The changes in teaching contents and methods of assessment

Year	Teaching hours of semester 1	Teaching hours of semester 2	Teaching Contents of semester 1	Teaching Contents of semester 2	Formative assessment	Summative assessment (achievement test)	Learning outcomes of University graduation
2013	80 periods	55 periods	From Unit 1 to Unit 14 of Lifeline textbook (Elementary)	From Unit 1 to Unit 6 of Lifeline textbook (Pre-intermediate)	Questions and Answers or Writing Test (Grammar or Reading exercise)	Writing Test (Grammar and Reading exercise)	
2017	80 periods	55 periods	Four skills and grammar/ vocabulary of KNLNNVN level 1	Four skills and grammar/ vocabulary of KNLNNVN level 2	Speaking Test/ Reading Test/ Listening Test or Writing Test	Writing Test (Objective test and Writing test)	EPT.2 of KNLNNVN

(b) Teachers of English are encouraged to use texts taken from journals, books and news for listening/speaking/reading and writing skills. The practices are designed by teachers were short answer

questions, gap-filling/identifications sentences/paragraphs, etc. that are identical to EPT.2 of VNFLPF or practice tests at A1 and A2 level.

The analysis of the official course documents indicated the official course documents were set before 2013 for semester 1 and 2 were not EPT.2 of VNFLPF or practice tests at A1 and A2 level. This shows that the impact of those on the teaching before 2013. Since 2014, a new trend has been seen: teachers of English have been encouraged to use a variety of authentic materials besides the official course documents. Thus, caution must be taken when interpreting the official course documents. This is also an issue that was mentioned in the interview with leaders and teachers.

Supplementary materials used by teachers

Leaders claimed that the formative assessment and semester examinations of English are similar to EPT.2 and CESOL tests (apart from the sub-writing of semester examinations). However, because of time limitation and mixed-big size class, one of four sub-tests (listening/speaking/reading/writing test) is applied for both formative assessment and semester examinations at HLU. The analysis indicated that a part of the semester examination focused on testing the mastery of grammar structures and vocabulary and that type of English test had stayed unchanged. There have been changes in the nature of the examination and the changes in question look undifferentiated to EPT.2 and CESOL item types and content. Hence, the interpretation must be that the semester examinations were shaped on the EPT.2 and CESOL examinations in the four sub-tests (listening/speaking/reading/writing test) as far as item types and content are concerned.

In short, results of the analysis of the supplementary materials practiced by teachers of English and students indicated they used various authentic materials that including commercial publications, journals, books and news for listening/speaking/reading and writing skills. They covered most Cambridge ESOL materials [23] (CESOL for short) that were available in Vietnam. Teachers and students did not use other kinds of materials (This is dealt with in the results of the interview and observation). The effects of CESOL tests were seen in the official course documents, but these materials were chosen after 2013. Teachers of English tended to use materials from CESOL sources to prepare students for semester examinations and EPT.2 examination. The analysis designates that other kinds of materials have no any influences on teachers and students. It must thus be deduced that EPT.2 and CESOL examinations have an impact on the choice of materials for teachers of English and students in classroom activities.

Results of Questionaires

There were differences between the respons of Teachers and Students on teaching listening/ speaking/ reading and writing activities in the classroom.

The responses of 715 students showed that the contents of their learning didn't focus on four skills (listening/ speaking/ reading and writing activities), whereas the responses of 14 teachers showed that the contents of their teaching focused on four (skills listening/ speaking/ reading and writing activities). The differences respons of Teachers and Students on teaching and learning activities in the classroom as the following Table 4.2.1, Table 4.2.2, Table 4.2.3 and Table 4.2.4

Table 4.2.1 Differences between the respons of Teachers and Students on teaching Listening activities in the classroom

Variables			Resp-types	Mean Difference	Std. Error Difference	T-test	df	2-Tailed Probability
While-Listening	Topics	Daily life (C2.1.5.1)	S	-1.4763	.1529	-9.654	12.320	.000*
			T					
	Participant organization	Pair work (C2.2.B.9.1)	S	-.8830	.2333	-3.785	11.869	.003*
			T					

*Note: S = Student; T = Teacher; * significant at p<0.05*

The survey data in Table 4.2.1 suggested that teachers and students varied little in terms of listening writing practices for English study. Only 2 (C2.1.5.1 of Topic and C2.2.B.9.1 of Participant organization) of the 46 survey items had statistically significant differences (p<.05) in terms of listening practices for English study.The differences in Table 4.2.1, though small, may be attributed to the influence of EPT.2 on listening teaching and learning in the classroom.

Table 4.2.2 Differences between the respons of Teachers and Students on teaching Speaking activities in the classroom

Variables			Resp-types	Mean Difference	Std. Error Difference	T-test	df	2-Tailed Probability
While-Speaking activities	Topics	Work and jobs (C2.1.4.2)	S	-1.2510	.1981	-6.314	12.218	.000
			T					
		Daily life (C2.1.5.2)	S	-1.4218	.1564	-9.091	13.481	.000
			T					
		Health and weather (C2.1.8.2)	S	-.9315	.2008	-4.639	12.461	.001
			T					
		Likes and Dislikes (C2.1.12.2)	S	-.7914	.1474	-5.368	12.734	.000
			T					
	Text	Filling a form (C2.2.A.15.2)	S	-.6028	.1483	-4.064	13.043	.001
			T					
	Participation Organization	Pair work (C2.2.B.9.2)	S	-1.4968	.1490	-10.046	13.279	.000
			T					
		Open questions (C2.2.B.14.2)	S	-1.4176	.2316	-6.120	11.533	.000
			T					

*Note: S = Student; T = Teacher; * significant at p<0.05*

The survey data in Table 4.2.2 showed that the respones of 14 teachers differed from the respones of 715 students in terms of speaking practices for English study.

There were 7 (C2.1.4.2/C2.1.5.2/C2.1.8.2/C2.1.12.2 of Topics and C2.2.A.15.2 of Text and C2.2.B.9.2/ C2.2.B.14.2 of Participant organization and) of the 46 survey items had statistically remarkable differences (p<.05) in terms of listening practices for English study.The differences in Table 4.2.2, though small, may be attributed to the influence of EPT.2 on speaking teaching and learning in the classroom.

Table 4.2.3 Differences between the respons of Teachers and Students on teaching Reading activities in the classroom

Variables			Resp-types	Mean Difference	Std. Error Difference	T-test	df	2-Tailed Probability
While- Reading activities	Topics	Work and jobs (C2.1.4.3)	S	-1.4875	.1787	-8.325	12.207	.000
			T					
	Texts	Letters and email (C2.2.A.6.3)	S	-1.2272	.1838	-6.676	12.112	.000
			T					
		Filling a form (C2.2.A.15.3)	S	-.3792	.1592	-2.383	13.661	.032
			T					
	Question types	Multiple choice cloze (C2.2.B.3.3)	S	-1.1566	.1500	-7.713	13.629	.000
			T					
		Open questions (C2.2.B.5.3)	S	-1.1292	.1378	-8.197	13.634	.000
			T					

*Note: S = Student; T = Teacher; * significant at p<0.05*

Table 4.2.3 presented teachers and students varied little in terms of reading practices for English study. There were 5 (C2.1.4.3 of Topic and C2.2.A.6.3/ C2.2.A.15.3 of Texts and C2.2.B.3.3/ C2.2.B.5.3 of Question types) of the 46 survey items had statistically noteworthy differences (p<.05) in terms of reading practices for English study.The differences in Table 4.2.3, though small, may be attributed to the impacts of EPT.2 on reading teaching and learning in the classroom.

Table 4.2.4 Differences between the respons of Teachers and Students on teaching Writing activities in the classroom

Variables			Resp-types	Mean Difference	Std. Error Difference	T-test	df	2-Tailed Probability
While – Writing activities	Topics	Work and jobs (C2.1.4.4)	S	-1.2887	.1787	-7.210	12.226	.000
			T					
		Transport (C2.1.6.4)	S	-.7251	.1459	-4.968	12.228	.000
			T					
	Texts	Letters and email (C2.2.A.6.4)	S	-1.2961	.1534	-8.448	12.484	.000
			T					

*Note: S = Student; T = Teacher; * significant at p<0.05*

It can be seen in Table 4.2.4 that the respones of 14 teachers differed from the respones of 715 students in terms of writing practices for English study. Only 3 (C2.1.4.4/C2.1.6.4 of Topics and C2.2.A.6.4 of Text) of the 46 survey items had statistically significant differences (p<.05) in terms of writing practices for English study. The differences in Table 4.2.4, though small, may be attributed to the influences of EPT.2 on writing teaching and learning in the classroom.

In short, this study presented an alternative approach to integrate the information gained from the responses of 14 teachers and 715 students to a series of questionnaire items with item observed by the researcher from direct observations to construct relevant variables. Each item of questionnaires was an evidence of the impact of EPT.2. All the items were designed onto the same scale to measure relevant dimensions of EPT.2 impact on the methods and contents of teaching, learning and curriculum. The small differences of survey data in Table 4.2.1, Table 4.2.2, Table 4.2.3 and Table 4.2.4 may be attributed to the influences of EPT.2 on listening teaching and learning in the classroom and thus, this finding is consistent with the view of Alderson and Wall (1993) and Cheng (1999 and 2004).

Results of observations

Because of the small differences of survey data in part 4.2 (Results of Questionaires), I decided to observe 10 teachers to see whether washback of VNFLPF and EPT.2 existed in their classrooms.

Round 1

The researcher observed ten classes of English A2 in semester 2. Two teachers of them and their students agreed to be videotaped. Ten teachers were female, with ten years of experience. Ten teachers and their students used materials from CESOL type. Textbooks were English File third edition A2 (Oxford, 2012). The supplementary materials are Cambridge Key English Test 1, 2 and English Grammar in Use of Murphy (2011) [24]. Ten teachers focused on four skills, grammar and vocabulary during Round 1 (listening: 9.2%/ speaking: 15.6%/ reading: 13.6%/ writing:10.2%/ grammar: 17.8%/ Vocabulary: 26%/; Pronunciation: 4%). Students worked in pair or group-work and made presentations and then ten teachers corrected their errors sometimes (3.8%). Because of time limitation, not all students could have a chance to speak English. Ten teachers used authentic materials (53.67%). The others were designed by themselves.

Round 2

The researcher observed ten classes of English A2 with 30 English lessons in Round 2. Ten teachers were female, with ten years of experience. Ten teachers and their students used materials from CESOL type. They didn't use textbooks. The supplementary materials are Cambridge Key English Test 1, 2 and English Grammar in Use of Murphy (2011). Ten teachers focused on practicing reading and writing (46.67%) and written test (24%) during Round 2. Students worked in pair or group-work and made presentations and then ten teachers corrected their errors sometimes (7.47 and 1.13%). Because of time limitation, not all students could have a chance to speak English.

In short, the result of observations corresponded to the responses of teachers and leaders. Ten teachers used a variety of materials from CESOL type. These materials are in line with the EPT.2 of VNFLPF 's approach. The methodology of ten teachers was communicative approach. It was hard to define whether the EFL teaching methodology was influenced by EPT.2 of VNFLPF 's approach or by the methodology of the used materials. However, this is an indication of EPT.2 of VNFLPF 's existence on EFL teaching. This is relevant to result of Alderson and Hamp-Lyons (1996), Cheng (1999 and 2004) and Pan and Newfields (2012).

Results of Interviews

Informal conversational interviews were conducted with 9/14 teachers of English after four-classroom observations and in groups at the office during tea break. Semi-structured interviews were held with the Rector of HLU, Head of Training Department and two Vice Directors of Foreign Language Centre, Relevant results are presented below.

85.71% teachers of English had already obtained M.A. degrees. One of them got C1 and four other teachers had obtained M.A. degrees at universities in either Australia or the USA. However, all teachers experienced over 7 years of teaching EFL and thus they could understand the changes on the national and institutional policies on EFL teaching and learning between 2013 and 2014. 100% of teachers often collected materials of CESOL, EPT.2 and CESOL-type to use in class. They also asserted that there were many practice tests for EPT.2 and CESOL examinations. They reported that they had been using them because materials for CESOL-tests were included in the office course documents and therefore they did not design task for students. They also expressed that they wanted their students to be familiar with numerous text contents and types of the principle of the EPT.2 and CESOL-tests. This helps the indication that there is interaction between teaching and learning and that this is interaction is related to the washback of KNLNNVN and EPT.

From these comments, they may be inferred that there had been many more materials on the market that were designed to prepare for EPT.2 and Cambridge ESOL examinations. It could also be said that teachers reacted differently to the needs of the test and self-designing tasks were also a problem for inexperienced teachers.

The selection of supplementary materials in is an indicator of VNFLPF and EPT washback on the use of materials.

Some of teachers did not think that they taught to the tests, they claimed that they taught to expand student's English. Thus, teachers described that reveal the trend to advocate the EPT.2 and CESOL-tests. In addition, nearly 70% of teachers said that they change their teaching methods to demand the changes of formative assessment and semester exams.

According to the Rector of HLU, the number of students admitted to HLU was increasing to meet the demands of society, and society demanded a high quality of training outcomes, particularly English proficiency of students. That was why the assessment of EFL learning outcomes at HLU must be innovated to meet the necessities of society. The Rector asserted that he wanted to maintain the institutional policies on English teaching according to VNFLPF next years because of its useful.

Head of Training department and two Vice Directors of Foreign Language Centre asserted that the semester exams of English were shaped on EPT.2 and CESOL-tests and that they were EPT.2-type, except for writing sub-test and the score scheme. Furthermore, teachers of English were acquainted with EPT.2 and CESOL-tests and they understood that the semester exams of EFL were shaped on EPT.2 and CESOL-tests. They believed that their tests were standardized because their tests were designed on EPT.2 and CESOL-tests. Therefore, the semester exams of EFL positively influenced curriculum designers, the EFL teaching and learning at HLU. Their answers illustrated that the tests in use are evidence of EPT.2 of VNFLPF washback.

In short, the responses of teachers and leaders revealed that there were EPT.2 and VNFLPF washback on what teachers used and on semester exams or in other words, teaching contents. Teachers agreed that formative assessment and semester exams corresponded to one of EPT.2 sub-tests. Nonetheless, few teachers supported that there was evidence of content washback on what they used. Accordingly, VNFLPF and EPT.2 have various types of washback on some teachers and learners than on other teachers and learners. This is relevant to result of Alderson and Hamp-Lyons (1996).

V. Conclusions And Suggestions

The insights from the findings show that VNFLPF and EPT.2 influence both positively and negatively the institutional policies on curriculum, the assessment of EFL learning outcomes and EFL teaching and learning in classrooms activities at HLU.

VNFLPF and EPT.2 have been considered one of the dominant determiners of what happens in classrooms that influence EFL teaching activities at HLU. The influences have been classified directly and indirectly, either positively or negatively. The curriculum, the official course documents, methods of assessment, methods of teaching and supplementary materials are innovated by the positive influences of tests. However, some inexperienced teachers did not design the tasks for students but relied on the available materials in the market that were related to negative washback.

Accordingly, the findings suggest that Ministry of Education and Training should issue a set of pre-constructed English tests that is modeled on EPT or Cambridge ESOL tests and then all schools would draw from this set to design their own version. In addition, teachers should be trained in educational evaluation and measurement that help them to design tasks or tests for their own students. This should help Vietnamese policy-makers, educators, and test writers, test users, teachers of English to prepare favorable conditions for enhancing

the beneficial washback of VNFLPF and EPT.2. The findings have contributed to the knowledge of the nature of washback and opened a new view to identify their different levels of washback effects.