

**VIETNAM NATIONAL UNIVERSITY, HANOI
UNIVERSITY OF LANGUAGES AND INTERNATIONAL STUDIES
FACULTY OF ENGLISH LANGUAGE TEACHER EDUCATION**

GRADUATION PAPER

**WASHBACK OF THE PROFICIENCY
ASSESSMENT AS EXIT REQUIREMENT
ON NON-ENGLISH MAJOR STUDENTS AT A LAW
SCHOOL**

Supervisor: Dương Thu Mai, Ph.D

Student: Bùi Thanh Hùng

Course: QH2014.F1.E3

HÀ NỘI – 2018

**ĐẠI HỌC QUỐC GIA HÀ NỘI
TRƯỜNG ĐẠI HỌC NGOẠI NGỮ
KHOA SƯ PHẠM TIẾNG ANH**

KHÓA LUẬN TỐT NGHIỆP

**TÁC ĐỘNG CỦA BÀI ĐÁNH GIÁ NĂNG LỰC
TIẾNG ANH CHUẨN ĐẦU RA ĐỐI VỚI SINH VIÊN
KHÔNG CHUYÊN KHOA LUẬT**

Giáo viên hướng dẫn: Dương Thu Mai, Ph.D

Sinh viên: Bùi Thanh Hương

Khóa: QH2014.F1.E3

HÀ NỘI – 2018

DECLARATION

I hereby state that I: Bùi Thanh Hoàng, being a candidate for the degree of Bachelor of Arts (English teaching Education) accept the requirements of the College relating to the retention and use of Bachelor's Graduation Paper deposited in the library.

In terms of these conditions, I agree that the origin of my paper deposited in the library should be accessible for the purposes of study and research, in accordance with the normal conditions established by the librarian for the care, loan or reproduction of the paper.

Signature

.....

May 2018

ACKNOWLEDGEMENTS

Firstly, I owe a debt of gratitude to my supportive and inspiring supervisor, Ms. Duong Thu Mai, Ph.D., who has offered me unceasing assistance to my study. I am sincerely thankful for her expert guidance, great encouragement and valuable advice while my research was carried out.

I would like to express my special thanks to students at the School of Law for their willingness to participate in the research. Without them, this study could not be fulfilled.

I am deeply thankful to my lecturers, my friends and my classmates for their constructive comments, helpful suggestions as well as constant stimulation. Their help plays a critical role for the completion of this study.

Last but not least, I would like to show my deep gratitude to my family, who have always inspire me to complete this study.

Hanoi, May 2018

Bùi Thanh Hùng

ABSTRACT

With the introduction of the Project 2020, students at most universities in Vietnam have to pass the graduation proficiency assessment. The understandings of the test effects on learning and teaching, or washback, are essential for the teachers, students and test administrators to gain desirable test outcomes. However, little is known about test's washback, particularly washback on students in Vietnam. This study, hence, investigate the perception of students of a graduation proficiency assessment and the **washback** of the test on students's test preparation strategies through questionnaires and interviews. The results show that students have positive perceptions towards the graduation proficiency assessment but the test exerts little effect on their test preparation. Although students of different years share the optimistic attitudes toward the test, they tend to prepare for the test differently. While the seniors study more test-like materials, the freshmen are likely to study general English tasks.

Table of contents

ACKNOWLEDGEMENTS	i
ABSTRACT	ii
LIST OF ABBREVIATIONS	v
LIST OF FIGURES AND TABLES	vi
CHAPTER I: INTRODUCTION.....	1
1. Statement of research problem & questions	1
2. Scope of research	2
3. Significance.....	2
4. Design of the study	2
CHAPTER 2: LITERATURE REVIEW	4
2.1. Issues in language assessment.....	4
2.2. Washback in language testing and assessment	4
2.2.1. The definitions of washback.....	4
2.2.2. The nature of washback.....	6
2.2.3. Theoretical framework of washback	7
2.3. English language proficiency tests.....	11
2.3.1. Communicative language proficiency	11
2.3.2. Major proficiency tests	11
CHAPTER 3: METHODOLOGY.....	16
3.1. Context	16
3.2. Research questions	16
3.3. Participants and selection of participants.....	17
3.4. Data collection method	18
3.4.1. Questionnaire.....	18
3.4.2. Interview	20
3.5. Data collection procedure	21
3.6. Data analysis	23
CHAPTER 4: RESULTS AND DISCUSSION	26
4.1. Students' attitude toward the graduation proficiency assessment (GPA)....	26

4.1.1. Students’ decision to take the GPA	26
4.1.2. Whether the test would motivate students to improve their English.....	27
4.1.3. Whether the test would raise students’ English proficiency	29
4.1.4. Whether the test would measure students’ proficiency accurately	30
4.1.5. Whether the test’s results would be useful in supporting job application	32
4.1.6. Test anxiety.....	34
4.2. GPA washback on students test preparation for the graduation proficiency assessment	36
4.2.1. The general English studies	36
4.2.2. Test-specific English activities	38
PART V: CONCLUSION	48
5.1. Summary of the findings and discussion	48
5.1.1. Students attitude toward the test.....	48
5.1.2. Washback on students’ test preparation	49
5.2. Conclusion and implication	50
5.3. Recommendations	50
5.4. Limitations of the study	51
5.5. Suggestions for further research	51
REFERENCES	52
APPENDIX	56
Appendix 1	56
Appendix 2.....	61
Appendix 3.....	66
Appendix 4.....	68
Appendix 5.....	70

LIST OF ABBREVIATIONS

FYSY	First year and second
GPA	Graduation Proficiency Assessment
MCQs	Multiple choice questions
MoET	Ministry of Education and Training
SPSS	Statistical Package for the Social Sciences
TOEFL	Test Of English as a Foreign Language
TYLY	Third year and last year
VNUH	Vietnam National University Hanoi
VSTEP	Vietnam Standardized Testing of English Proficiency

LIST OF FIGURES AND TABLES

Figure 2.1. A basic model of washback (Hughes, 1993)	8
Figure 2.2. A washback model of students' learning (Shih, 2007, p.151).....	10
Figure 3.3. Interview participants	22
Figure 4.1. Reasons to take the GPA.....	27
Figure 4.2. Students' perception toward test motivation	28
Figure 4.3. Students' perception toward test function of raising English proficiency	29
Figure 4.4. Students' perception of the GPA's accuracy	31
Figure 4.5. Students' perception toward the test usefulness	33
Figure 4.6. Respondents' test anxiety	35
Figure 4.7. Students general English learning activities	37
Figure 4.11. Time to prepare before the test	40
Figure 4.14. Test preparation strategies	43
Figure 4.15. Time for each test preparation activity	45
Figure 4.16. Difference between two groups	46
Table 2.3. The structure of the TOEFL iBT Test.....	12
Table 2.4. VSTEP Background (Carr et al., 2016).....	13
Table 3.2. Questionnaire Items.....	20
Table 3.4. Reliability Statistics part 1	24
Table 3.5. Reliability Statistics part 2	24
Table 3.6. Reliability Statistics part 3	24
Table 4.8. Number of General learning activities	37
Table 4.9. Numbers of hours per week for General English activities	38
Table 4.12. Time for test preparation of respondents per week.....	41
Table 4.13. Number of learning strategies taken by respondents	42

CHAPTER I: INTRODUCTION

1. Statement of research problem & questions

The impacts of language assessment, in other words, test washback, have seen an increased interest in the last few decades (Alderson & Wall, 1993; Bailey, 1996; Hughes, 1989; Caine, 2005). The research contexts spread from Canada, Brazil, Germany, and Greece to Japan, China, and Hong Kong. A large number of studies have emerged, especially on large scale tests such as TOEFL and IELTS (Alderson & Hamp-Lyons, 1996 and Green, 2007), Hong Kong Certificate of Education Exam in English (HKCEE) (Cheng, 2004), General English Proficiency Test (Pan & Newsfields, 2012). The findings of the degree of washback, however, vary overtime; while some scholars concluded that there was almost no washback in their studies (Watanabe, 2014; Shih, 2007), others stated the strong effect of washback (Ferman, 2004). Thus, it can be indicated that the washback might be different in each context.

In context of Vietnam, in recent years, the status of foreign languages is considerably heightened thanks to the implementation of National foreign language 2020 project, which requires that “by 2020, most Vietnamese students graduating from secondary, vocational schools, colleges and universities will be able to use a foreign language confidently in their daily communication, their study and work (...)” (Prime Minister, 2008). Therefore, universities and colleges have to select a Graduation Proficiency Assessment (GPA) for their students. One of the tests used for university exit is known as “*Vietnam standardized test of English proficiency (VSTEP), a high-stake test which is likely to exert certain impact on test users (e.g. students, teachers and practitioners)*”. However, there are a few studies of GPA or VSTEP washback on students, particularly the non-English majored ones.

To fill such gap, this study aims to investigate the washback on non-English majored students’ preparation strategies for the graduation proficiency assessment (GPA) and their attitudes toward the test in VNUH, particularly at the School of Law.

In brief, this study will address the following questions:

1. What are students' perceptions toward the English graduation proficiency assessment?
2. To what extent does the graduation proficiency assessment influence students' preparation strategies(from students' perspective)?
3. Are there differences in the test preparation strategies of students in different years?

2. Scope of research

This research targets students at the School of Law, Vietnam National University, Hanoi (VNUH). To graduate from this school, all students are required to pass the GPA level of B1 or B2; or in other words, they have to pass the level 3 or level 4 of Vietnamese Standardized Test of English Proficiency (VSTEP). The alternatives namely TOEFL, TOEIC, Cambridge Test, or IELTS certificates are also accepted. Moreover, students can make decision to take any examination as long as they have an equivalent level 3 or 4 VSTEP certificate which satisfies the language graduation requirement.

For time limited, this study will only look at the washback on students' preparation for the test and their attitude toward the GPA.

3. Significance

Despite being conducted at a small scale, not to be generalized, the study would partly fill the literature gap of washback on students. By targeting non-English major students, the research may raise their awareness of different test preparation strategies as a good source for better learning. Simultaneously, teachers, practitioners and other stake holders might have a better understanding of their students' attitude and learning in order to enhance the quality of the tests, teaching and learning activities. This study would also be useful for other researchers concerning language assessment, particularly washback area.

4. Design of the study

This study is divided into five chapters:

Chapter 1: Introduction – presents the rationale, aims research questions, significance, scope, and design of the study

Chapter 2: Literature review – discusses the theoretical framework and studies of washback and English language proficiency tests

Chapter 3: Methodology – demonstrates the context and research design of the study

Chapter 4: Findings and discussion – shows the comprehensive of data and a discussion on the findings of the study

Chapter 5: Conclusion – provides a summary of the findings, conclusion, recommendation, limitations, and future direction for the study.

CHAPTER 2: LITERATURE REVIEW

2.1. Issues in language assessment

Historically, language testing has evolved and expanded over 30 years. The 1980s witnessed a shift from language tests which focused on discrete-point format to the communicative language testing; besides, the field also expanded to Second Language Acquisition areas. The next decade carried on the expansions in numerous areas namely “research methodology”, “practical advances”, “factors that affect performance on language test”, authenticity, “ethical issues and consequences of test use.” (Bachman, 2000, p.4). Of all those issues, the study of washback, a facet of test consequences, until now, has still been discussed worldwide for its importance. According to some researchers, the test impact could “govern and determine people’s future education” (Shohamy, 2001 as cited in Brown & Abeywickrama, 2010, p.22). Due to the considerable power and authority of the tests, many policy-makers may use them to administer the educational system or promote particular behaviors (Shohamy et al., 1996).

2.2. Washback in language testing and assessment

2.2.1. The definitions of washback

Testing has been utilized in education and employment area for such a long time for a multitude of purposes namely measuring test-taker’s proficiency, playing the gatekeepers’ roles, or motivating students. On account of the powerful role of test in different fields, there is a notion that testing has an influence on learning and teaching. Before 1990s, the common terms in language testing to refer to the belief about the testing and teaching/ learning relationship were “curricular alignment” (Linn, 1983), which refers to the relationship of the testing content and the designed curriculum; “measurement-driven instruction” (Popham, 1987), which holds the view that “testing should drive curriculum and thereby teaching/ learning”; “systemic validity” (Frederiksen & Collins, 1989), which refers to changes in curriculum and instruction to “foster the development of the cognitive skills that the test is designed to measure (Frederiksen & Collins, 1989, p.27). In 1993, for the first time, Alderson and Wall official introduced the

term “washback” with proven evidence from empirical studies. Since then, more attention has been paid to washback with different definitions.

Washback or backwash in language assessment is generally known as “the effect of testing on teaching and learning” (Hughes, 1989, p.1). Despite the prevalence of washback in the applied linguistics field, the term is hardly seen in dictionaries. However, some dictionaries such as the New Webster’s Comprehensive Dictionary and the Collins Cobuild Dictionary can be found including backwash which is defined as “the unwelcome repercussions of some social action” and “unpleasant after-effect of an event or situation” respectively. These definitions consist of an interestingly negative connotation which refers to the adverse relationship between testing and teaching/ learning discussed later. (Cheng et al., 2004). Alderson & Wall (1993) assume that washback renders “teachers and learners to do things they would not necessarily otherwise do” (p.117). Messick (1996) refers to washback as “the extent to which the introduction and the use of a test influences language and teachers to do things they would not otherwise do that promote or inhibit language learning” (p.241). “The connection between testing and learning” is how Shohamy et al. (1996) defines the term.

In terms of scope, washback is divided into narrow and broad scope. In the narrow scope, washback is more frequently referred to the effect of the test on teaching and learning. Both Alderson and Wall (1993), Bailey (1996) and Messick (1996) agree that “teachers and learners do things *they would not necessarily otherwise do* because of the test (Alderson and Wall, 1993, p. 117). In the broad scope, washback is just considered as a facet of consequential validity which “encompasses all the consequences of the test”, covering the concerns of its accurate intended-criteria measurement, its influence on the test preparation; and its “social consequences of a test’s interpretation and use” (Brown & Abeywickrama, 2010, p.34). Bachman and Palmer (1996) utilizes the term **impact** to refer to consequential validity, possibly “more broadly encompassing the many consequences of the assessment prior and after a test administration” (p.34). Therefore, washback is considered as one dimension of impact (Hamp-Lyons, 1997). For illustration, Bachman and Palmer (1996)

consider test impact into two levels: micro- and macro- levels. The impacts on individuals namely test-takers or teachers are at the micro level. At the macro level, the test can affect the whole educational system and society.

In this study, the interpretation of “washback” in narrow scope will be adopted: washback at the micro level which is the effects of the test on individuals. More specifically, the effect of VSTEP on students in the aspects of test preparation and attitudes.

2.2.2. The nature of washback

Regarding the nature of washback of a test, most studies mention the dimension of value such as positive or negative. However, according to Watanabe (2000), there are four other dimensions namely specificity (*general or specific*), intensity (*strong or weak*), length (*long or short*) and intentionality (*intended or unintended*) (Cheng at al., 2004, p.20). In this minor thesis, only the value and specificity of washback are discussed.

2.2.2.1. Positive versus negative

Alderson & Wall (1993) considers washback as a neutral term which could be beneficial or negative. The positive washback would have beneficial impacts such as promoting teaching learning activities, encouraging positive attitudes toward the test and improving motivation (Alderson & Wall, 1993). Shohamy (1993) also states that the test has positive washback if it forces students to learn more such as listening more carefully or taking the study more seriously. On the other hand, negative washback refers to the test anxiety and the fear of poor performance; hence, focus on excessively on the skills tested (Tsagari, 2011). Similarly, teachers might be afraid of the test results that they may just focus on the test or narrow the curriculums. Another feature of negative washback is that students review their notebooks and find books which are related to the test (Damankesh & Babaii, 2015). In short, the positive washback encourages the derisable changes while the negative washback bring undesirable ones.

2.2.2.2. General versus specific

This dimension of washback indicates that the washback may be *general or specific*. The general impact may be produced by any tests (i.e. students study harder), whereas, if the washback is specific, it can be seen in only one aspect or a type of test. Didi & Ridha (2011) state that the English national Examination (ENE) has specific washback because the teachers focused more on reading skills which were greater part of the test rather than communication skills.

2.2.3. Theoretical framework of washback

This part examines different models of washback of Hughes (1993), Bailey (1996), Alderson & Wall (1993) and Shih (2007). Firstly, a basic model of washback been early proposed by Hughes (1993) in which three components “participants”, “process” and “products” are differentiated. (Bailey, 1996, p.262)

According to Hughes, participants include students, classroom teachers, administrators, materials developers and publishers, 'all of whose perceptions and attitudes towards their work may be affected by a test' (1993: 2). Under process Hughes (1993: 2) includes 'any actions taken by the participants which may contribute to the process of learning'. Such processes include materials development, syllabus design, changes in teaching methodology, the use of learning and/or test-taking strategies, etc. Finally, product refers to 'what is learned (facts, skills, etc.) and the quality of the learning (fluency, etc.)' (1993: 2)

(as cited in Bailey, 1996, p.262)

Hughes (1993) considers that the nature of the test would influence the participants' thinking toward learning and teaching; hence, the learning and teaching behaviors (process) might be affected leading to an inversion in the outcomes (products).

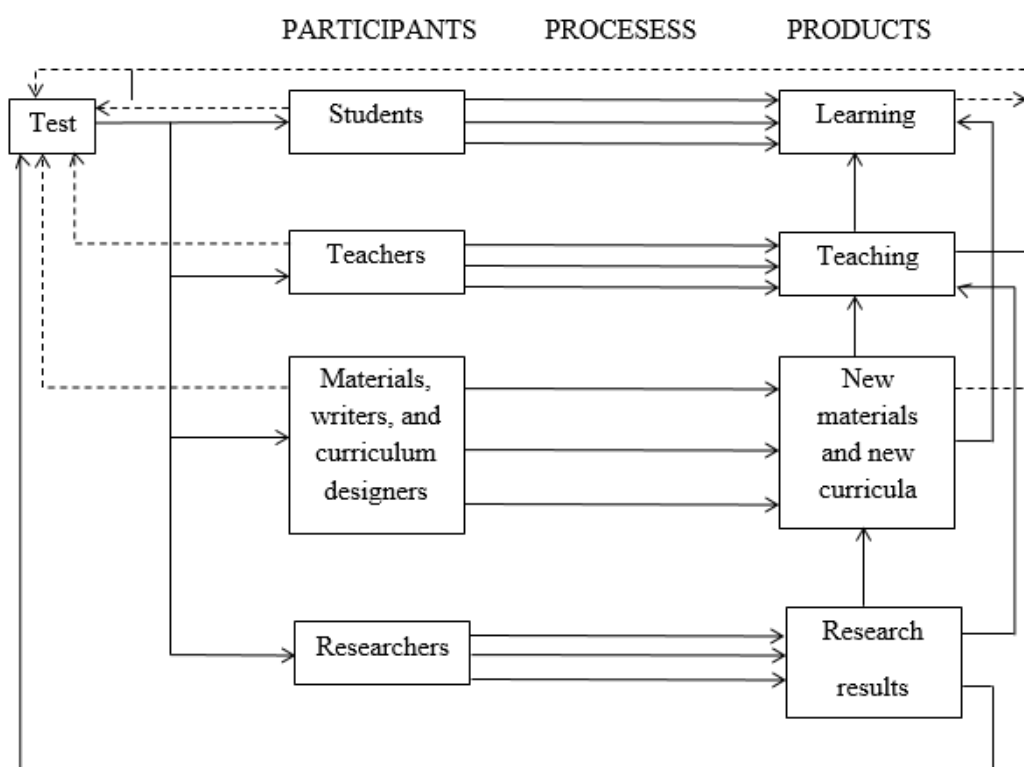


Figure 2.1. A basic model of washback (Hughes, 1993)

Despite showing the clear relation among the test and others elements namely participants, process and products, the model does not include different aspects affecting learning and teaching process.

Basing on model of Hughes (1993), especially participants and product, Bailey (1996) construct his model comprising of two parts 'washback to the learners' which refer to "the effects of test-derived information on test-takers and and "washback to the programme" which are washback to "teachers, administrators, curriculum developers, counselors, etc.,". (p.264)

Some examples of the "wash aback to the learners" are suggested by Bailey (1993) as following:

- 1) Practicing items similar in format to those on the test.
- 2) Studying vocabulary and grammar rules.
- 3) Participating in interactive language practice (e.g., target language conversations).
- 4) Reading widely in the target language.
- 5) Listening to non-interactive language (radio, television, etc.).
- 6) Applying test-taking strategies.

- 7) Enrolling in test-preparation courses.
 - 8) Requesting guidance in their studying and feedback on their performance.
 - 9) Enrolling in, requesting or demanding additional (unscheduled) test-preparation classes or tutorials (in addition to or in lieu of other language classes).
 - 10) Skipping language classes to study for the test.
- (Bailey, 1996, p.264)

Nevertheless, the model does not indicate how different individual is affected by the test. Alderson and Wall (1993)'s model which is considered more in details (Pan, 2014) proposed the "Washback Hypothesis" taking a different approach. There are "15 possible hypotheses regarding washback" as following:

- 1) A test will influence teaching.
 - 2) A test will influence learning.
 - 3) A test will influence what teachers teach; and
 - 4) A test will influence how teachers teach; and therefore, by extension from (2) above:
 - 6) A test will influence how learners learn.
 - 7) A test will influence the rate and sequence of teaching; and
 - 8) A test will influence the rate and sequence of learning.
 - 9) A test will influence the degree and depth of teaching; and
 - 10) A test will influence the degree and depth of learning.
 - 11) A test will influence attitudes to the content, method, etc., of teaching and learning.
 - 12) Tests that have important consequences will have washback; and conversely
 - 13) Tests that do not have important consequences will have no washback.
 - 14) Tests will have washback on all learners and teachers.
 - 15) Tests will have washback effects for some learners and some teachers, but not for others.
- (Alderson and Wall, 1993, p. 120-121)

The washback on students are now demonstrated in more aspects namely how and what learners learn, the rate, sequence, degree, and depth of learning and the attitude to content and method of teaching.

Shih's study (2007) built a new and more detailed model on washback on students with a thought that the previous models are not suitable any more. The model shows the variety of test components which might exert impacts, the different aspects on students' learning and psychology and other factors such as extrinsic and intrinsic ones. As can be seen from the chart, washback of the test

on students learning includes content of learning, total time on learning, learning strategies, learning motivation and test anxiety. However, how each aspect of learning as mentioned is affected by the test is not demonstrated.

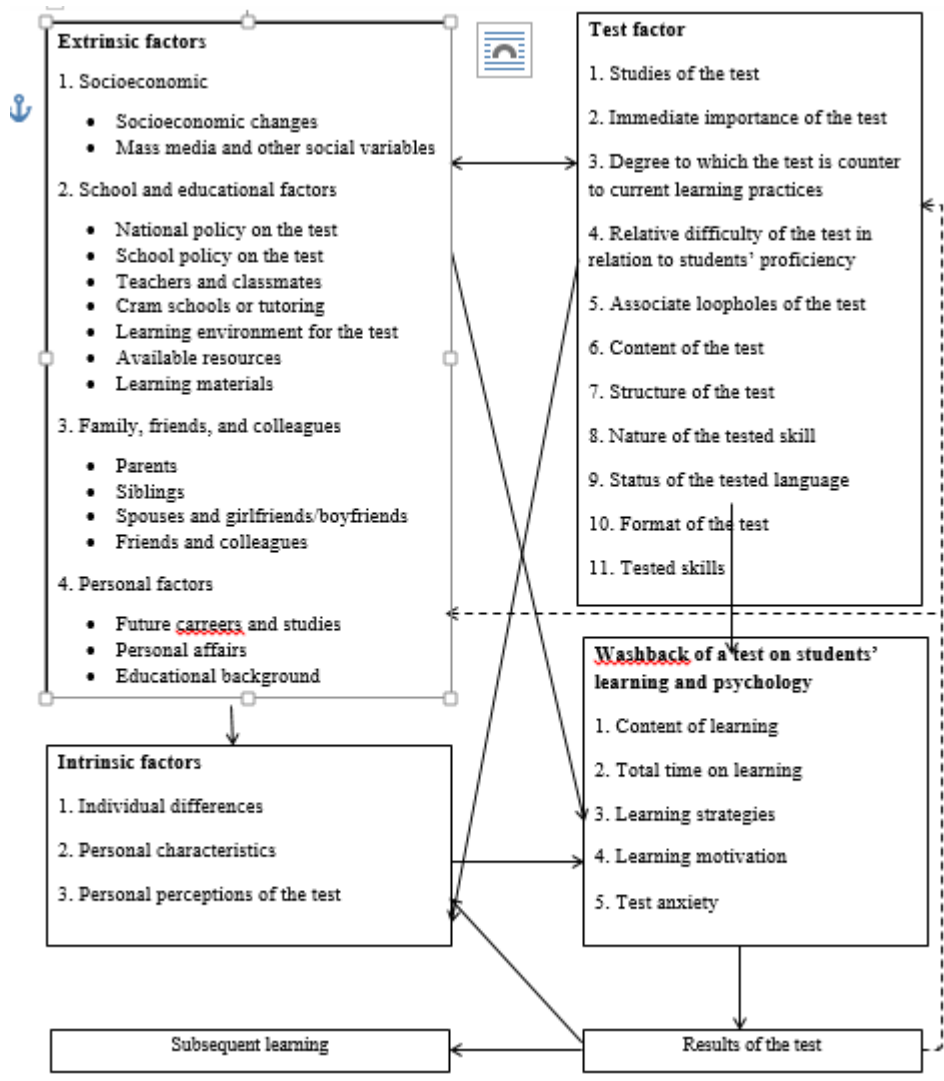


Figure 2.2. A washback model of students' learning (Shih, 2007, p.151)

Four frameworks of washback have been examined, while Hughes's (1993) model does not display different aspects of learning and teaching, Bailey (1996) adds some examples of washback on students. However, the framework of washback does not show how each individual is influenced. The Washback Hypothesis (Alderson & Wall, 1993) has predicted the influences of washback on both students and teachers on different aspects. Nevertheless, Shih (2007) has presented another model which is claimed to be more modern and focus more on students.

In this research, the model of Shih (2007) is adopted as it is more students focus. Five aspects of washback on students learning and psychology are also concurred by Ferman (2004) (cited in Cheng at al., 2004)

2.3. English language proficiency tests

To study washback, different examinations are investigated. As most of them are language proficiency tests, the following part concentrates explaining this test type and give some typical examples.

2.3.1. Communicative language proficiency

Communicative Language Testing (CLT) is an original approach toward language testing, which focuses on the authenticity and communicative purposes of the test. Proficiency test “is designed to measure people’s ability in a language, regardless of any training they may have had in that language” (Brown and Abeywickrama, 2010, p.11). Besides, it is also stressed that the test is not restricted to any course or skills. Therefore, a communicative language proficiency test examines the overall English of test-takers for communicative purposes.

Most communicative language proficiency tests have summative results which present a single score with 2-3 sub scores. These sub scores are less likely to give any feedback or serve diagnosis function; hence, students can only interpreting their scores by looking at the rubrics.

2.3.2. Major proficiency tests

There are hundreds of proficiency tests worldwide. The following section solely introduces an international test, TOEFL and a national examination, VSTEP which are now highly concerned in Vietnam.

TOEFL iBT

TOEFL is one of the most trustworthy English tests worldwide given via computer. The test is primarily sponsored by American Educational Testing Service (EST). The purpose of TOEFL iBT is to assess test-takers’ English proficiency whose English are not native language. The results are primarily

employed to decide the international students' ability of English in academic context.

The table below is the demonstration of the TOEFL iBT test structure. The test includes four sections Reading, Listening, Speaking and Writing allotted in about four hours. Each skill makes up 30 scores to set the total score of 120. Detailed information of each skill follows.

Table 2.3. The structure of the TOEFL iBT Test

Section	Number of items/ tasks	Testing time	Score Scale
Reading	36–70	60–100 minutes	0-30
Listening	34–51	60–90 minutes	0-30
Break		10 minutes	
Speaking	6 tasks	20 minutes	0-30
Writing	2 tasks	50 minutes	0-30
Total		Approximately 4 hours	0 - 120

As it can be seen from the format, the aim of communicative language is quite clear. Firstly, the test content resembles university-life context such as the lectures, conversations with friends and scholarly passage. In addition to language skills, other necessary skills of university are tested as well namely giving opinions, summarizing the passage, and analyzing the information.

VSTEP

VSTEP 3-5 is the first ever Vietnam Standardized Test of English Proficiency issued by University of Languages and International Studies (VNU), VNUH since May 2015. It is to serve the national language assessment purpose of the National Foreign Language Project 2020.

The test is aimed to measure Vietnamese adult's English ability according to the CEFR-based National level 3-5 Proficiency Scale. The result which is recognized nationwide can be used for gate-keeping function, placement decision or in guiding English-language instruction. In VNUH, students have to

achieve a certain level to be admitted to a program or to graduate from the university.

VSTEP is based on Bachman (2002) and Bachman and Palmer (2010) theory, and validation framework which is Common European Framework of Reference for Languages (CEFR), and simultaneously focus on test uses. To illustrate, it is constructed to match Vietnamese studying and working environment with some contents of Vietnamese culture, economy and community integrated the tasks.

VSTEP 3-5 measure level 3 to 5 from National 6-level Proficiency Scale, which is compatible with level B1, B2, C1 from CEFR. The details are demonstrated in the following table:

Table 2.4. VSTEP Background (Carr et al., 2016)

Score	Level	CEFR
0 – 3.5	Not rated	Not rated
4.0 – 5.5	3	B1
6.0 – 8.0	4	B2
8.5 - 10	<u>5</u>	C1

The test comprises of 4 section each assess a skill. Test-takers spend nearly 3 hours completing the test. Each skill makes up 10 points and the total score, which is on the scale 10, is the average of 4 skills. The test format is presented in the Appendix 4.

2.4. The studies of washback

2.4.1. The studies of washback in the world

There are numerous studies of washback on teachers worldwide; however, washback studies on students remains in shortage (Shih, 2007). Watanabe (2004) emphasized on the need of more work on learners. Although washback is various in different contexts and individuals, the researchers agree that the aspects of washback on students include behavioral aspects and altitudinal aspects. Various impacts on these aspects have been explored.

Shohamy et al.(1996) investigate the washback of two different tests namely ASL and EFL from perspectives of teachers, students and language inspectors. The results show two different pictures. While the ASL has almost no effect on the teaching and learning activities, test preparation and time allotment, the EFL has tremendous washback. Although the teachers show negative attitudes toward the two tests' quality, the tests are still in need for learning promotion.

Pan (2014) examines the washback of TOEIC & GEPT as graduation requirements by doing survey, asking question and observing two groups including non-exist requirement students and exist requirement students. When looking at the aspect of degree/ depth of learning, attitude toward methods of learning and some learners but not others being influenced" (Alderson & Wall, 1993), Pan divides the questionnaire into three parts types of test-preparation activities, type of language skill-building activities, viewpoint of the GEPT and TOEIC. The results show that although the performance and motivation of the exist requirement group higher, the learning methods of two group are not significantly different.

Damankesh & Babaii (2015)_investigate a highschool final examination and_explore the washback on students' test-taking and test-preparation strategies. The results show that the examination influences students' learning behaviors by getting them take certain types of test-taking and test-preparation strategies. The washback is partly negative as some strategies have adverse effect on students' creativity and inhibit their learning. On the other hand, the test has some slightly positive influences which foster the students' cognition and attention.

In 2017, Green investigates the effectiveness of IELTS preparation class on students' writing scores. Students had to take two writing tests prior and after the preparation course. In addition to the test instrument, there were two questionnaires to gain data of participants' different background and the "process and outcome variables." It was concluded that the test preparation brought no apparent benefit to students' test scores. Another study on TOEFL (Alderson &

Hamp-Lyons, 1996), however, brings an opposite result that the test had influence on both what and how teachers teach.

Different communicative proficiency tests have been studied all over the world from international levels such as IELTS and TOEFL, national levels such as ASL, EFL and GEPT to the local tests such as high school examination. The results are varied among the studies. At the same scale, while TOEFL had washback on teachers on both how and what they teach, the IELTS preparation did not have apparent washback. Therefore, the test washback might not be affected by the test nature only. In order to understand washback, specific contexts need investigating. In the context of Vietnam, nevertheless, there are only a few studies.

2.4.1. The studies of washback in Vietnam

In Vietnam, there are three studies of washback, only one of which is about the VSTEP, the newly designed test in Vietnam. The other two studies are international examinations such as TOEFL (Nguyen, 1997 & Barnes, 2016)

Nguyen (2017) studied the washback of VSTEP on first year students at ULIS, VNUH. The two aspects of behaviors and attitudes were examined by Likert scale questionnaires. It is found that the test had strong washback on learning content, learning methods and affective conditions.

However, no study of washback in Vietnam has focused on non-English major students.

Chapter summary

This chapter has reviewed the language testing context and given an insight view about washback definition, nature and different models. This study adopts the narrow definition of washback, studies washback value and specificity and applies the model of Shih (2007). Two typical language proficiency tests were introduced followed by a number of studies worldwide and in Vietnam. There are a multitude of washback research on teachers but few studies have studied the washback on students particularly in Vietnam context.

CHAPTER 3: METHODOLOGY

3.1. Context

In Vietnam, the status of English has recently heightened thanks to the Project 2020 by Prime Minister (2008). To achieve the goals that most Vietnamese when graduating from universities and colleges could use English in Dailey conversation, the VSTEP 5 levels was issued. Most university students have to pass certain levels of the test in order to graduate from schools. Other compatible international certificates such as IELTS, TOEFL, TOEIC are also accepted.

The School of Law, VNUH, also adopts the VSTEP as the graduation proficiency assessment (GPA). The VSTEP level 3 or the alternative international tests is the precondition for students to graduate from the school. Only fast-track students have to achieve the level 4. As the certificates are valid in two years, most students consider taking the test in the second half of their students' life. It should also be noted that these students can take the test at any time they find convenient.

3.2. Research questions

This study examines 3 research questions:

1. What are students' perceptions toward the English graduation proficiency assessment?

By answering this question, the researcher penetrates how students perceive the test and if they are negatively or positively influenced. Their perceptions towards the GPA may also explain their behaviors.

2. To what extent does the graduation proficiency assessment influence students' preparation strategies (from students' perspective)?

This question aims to investigate students' reflection/perspectives regarding their behavior during the preparation phase to see whether the GPA exerts washback on them, whether those impacts are strong or week, negative or positive, general or specific.

3. Are there differences in the test preparation strategies of students in different years?

The question is raised to explore whether the first-year students and senior study English differently for the GPA.

3.3. Participants and selection of participants

The population size is approximately 1400 students including around 700 first-year and second year students and roughly 700 third-year and fourth-year students at the School of Law, VNUH.

Participants are students at the School of Law, VNUH. They are not major in English nor studying English for Special Purposes. The School of Law offers three General English courses from level A1 to B1 for mainstream students and one extra B2 course for students in fast-track division. However, students can choose to take the courses or not at any stage of their university life. While English course taking is flexible, the graduation proficiency assessment is compulsory. Students at mainstream program have to reach B1 level by the time they graduate and it is particularly B2 level for fast-track students. The School identifies students' levels by VSTEP certificates or other international certificates namely TOEIC, IELTS and TOEFL. As a result, students have to take the tests before graduating. First-year and second year students are newbies to the School; thus, they still have a plenty of time ahead to prepare for the test. In comparison, the third-year and four-year students are those who have limited time for test preparation.

Regarding sampling design, the study involves stratification of the population before selecting the sample. The population is divided into two strata based on the school year with the characteristics briefly described above. The first group comprises of the first and second year students while the third and fourth year students are in the other stratum. The ratio of two groups is roughly 1:1. For each stratum, cluster sampling is used. The law students study in different classes in Hall P1 and P2 in G Building, hence, some classes will be selected randomly for data collection.

The sample size was 123 students considering the population size 1400 students, confidence level 95%, confidence interval ± 8 .

In 123 responses, 82.9% students (102 students) report to take or have taken the GPA at school while 17.1% consider to sit alternative proficiency tests namely IELTS and TOEIC. Those who do not take the GPA will be excluded from the latter half of the study which examines students' test preparation behavior. In 102 students responding to take the test, there are 57 first year and second year students making up 55.9% and 45 third and fourth year students accounting for 44.1%.

In summary, the sample is reasonably representative of the School of Law GPA test-taking students population.

Eight respondents from the sample will be invited for the interview after finishing the questionnaires. They are those who are willing to take part in the interview round.

3.4. Data collection method

Face-to-face questionnaires and semi-structured interview are employed in this study for their advantages. The study does not apply observation method as in the context of the School of Law, there is no compulsory English course and students have choice to study in any suitable classes.

3.4.1. Questionnaire

Four research questions are answered with a questionnaire, which is the main instrument for the study because of its significant benefits compared with other instruments. Firstly, by nature, questionnaires are practical to deal with a large number of respondents. With the sample of 123 participants, questionnaires are the most manageable and reasonable to distribute surveys, analyzes responses and draw a general conclusion as compared to one-on-one or group interview or observation. Secondly, the instrument is arguably cost-effective, as the researcher can quickly collect a substantial number of responses from participants.

The face-to-face questionnaire is chosen as its advantages in administration. Face-to-face questionnaire gives the researcher to meet participants in person to explain the purpose of the study as well as to answer any pop-up questions. Besides, a higher rate of response will be guaranteed in comparison with online and mail questionnaires.

The questionnaire is adapted from Stoneman (2005). Questions related to students' self-rating and perspective toward the test, respondents' out of class activities and test preparation are raised in the survey. The questionnaire comprises of four parts as following:

Part 1 asks for demographic information. Questions about the name, phone, and email are not numbered and students only need to leave information if they are willing to join the interview round. Question 1's purpose is to find out whether students take the GPA or not. All the questions done by non-GPA test-takers will be excluded. As the purpose of this paper is to find the washback of GPA, if students take other tests, it is unlikely that they have any influences. Question 2 collects students' school year which helps the researcher to divide the participants into two groups as presented above in order to answer research question 3.

Part 2 helps to collect washback on attitudinal aspects. There are 7 questions from question 3 to 9. Question 3 asks for the reasons why students choose GPA but not other tests. The next two questions are students' self-rating of their English proficiency and test anxiety which might explain students test preparation later. The next four questions seek for students' perspectives of GPA as a source of motivation, as a means to raise students' standard, as a tool to measure students English proficiency correctly and a support for job application.

Part 3 includes 2 questions to find out students' learning activities outside classes; to be more specific whether students learn English by watching the TV program, listening to music, or speaking with native speakers. The second question aims at examining the average time students spend on doing those activities weekly

Part 4: includes 4 questions which are to investigate washback on students' behavioral and attitudinal aspects.

Table 3.2. Questionnaire Items

Part	Content	Item
Part 1: Background information	- demographic information (name, phone, and email) - GPA test-taking decision - students' school year	Q1 Q2
Part 2 Attitudinal aspects	- reasons to take the GPA (believes toward the test) - self-rating proficiency - respondents' level of test anxiety Perspective toward the test as: - as source of motivation for students to improve their English - as a means to help raise students' English standard. - as a tool to measure students' English proficiency - the usefulness of the test result in supporting job applications	Q3 Q4 Q5 Q6 Q7 Q8 Q9
Part 3 Behavioral aspects (learning activities outside classes)	- outside class activities - time spent on those activities per week	Q10 Q11
Part 4 Behavioral aspects (test preparation)	- test preparation status - time spent preparing for the test - types of test preparation activities - the content of materials	Q12 Q13 Q14 Q15

3.4.2. Interview

As participants cannot be directly observed, interview is a reasonable method to collect participants' historical information (Creswell, 2014) such as the process that students prepare for the test.

Semi-structured interview is also utilized because there are disadvantages in the use of questionnaires. Firstly, it is argued to be insufficient in

understanding some information (i.e. behaviors, feelings). Secondly, the questionnaire includes limited questions developed from a fixed scheme, so they may have some missing important information. In order to overcome these weaknesses, semi-structured interview is put into use.

The interview questions are aimed not to draw generalization but to provide essentially qualitative explanations for certain findings in the questionnaire. The interview scheme is divided into 3 parts. Part 1 – identify students test taking status and the reasons they choose this examination. Part 2 asks questions about students’ perspectives toward the GPA and why they hold those opinions so as to answer Research question 1. Students perspectives toward test importance, test results, test motivation, test accuracy and test anxiety, as well as their motives, are collected in this part. Last but not least, behavioral aspects of washback such as time allotment, learning strategies, content, and materials are covered in part 3.

3.5. Data collection procedure

Step 1: Designing the survey questionnaire and interview questions

Both questionnaire and interview questions are designed based on washback theory on students’ test preparation and modified from previous studies. While the questionnaire depends heavily on Stoneman’s paper (2005), the question list is adapted from both Stoneman (2005) and Shih (2007).

As the participants are not major English, the questionnaire and interview questions are translated into Vietnamese. By doing that, not only is misunderstanding, or misinterpreting shunned, but also the respondents are able to express themselves more precisely and thoughtfully.

Step 2: Piloting and revising instruments

The questionnaire was piloted by 7 students as presented in chart below.

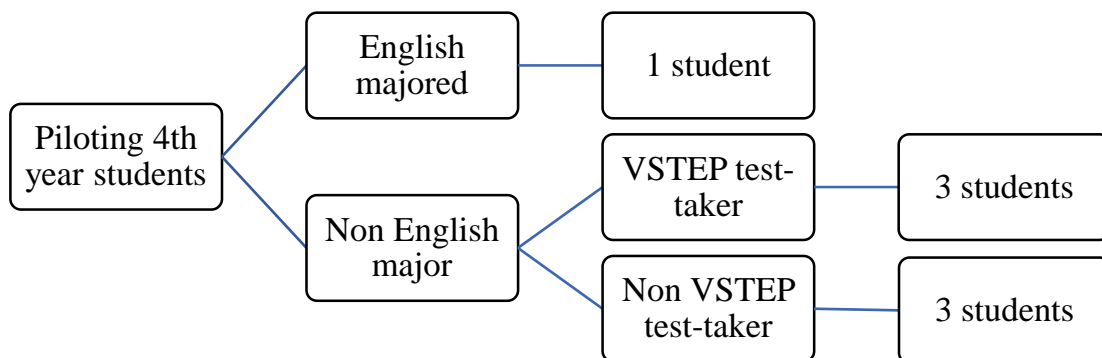


Figure 3.3. Interview participants

The unnatural word choice, ambiguous questions and overlapping options were detected. Many students also complained that the questionnaire was lengthy and included ample open questions. Taking the feedbacks seriously, the researcher refined the questionnaire. Firstly, some questions namely question 1, 6 and 15 were reworded so it sounded more Vietnamese. Then, some items in question 10 were merged. Questions asking “when did you first access VSTEP website/ go to the English center/ prepare for VSTEP” were omitted for students could not remember the exact time.

Step 3: Distributing questionnaires to the selected class and ask the participants to fill in their responses.

Survey:

The researcher gains timetable of the School of Law in the second term of the school year 2017-2018 and gets to random classes in the break to distribute questionnaire. It was effortless to locate the first and second year students while the classes including the juniors and seniors are tougher to find. Therefore, the research came up with an online questionnaire to expand the number of responses from the latter group.

While delivering questionnaire, the researcher explains the objectives of the study and informs the following interview section so the respondents could fill in their phone number and email if they consented to take part in. As the researcher had to approach different classes during break time and one or two times the lecturers came in before the students finished the questionnaire, the research was not with the participants the whole time to

answer pop-up questions. However, some queries were still clarified while the researcher was in the class.

Interview

20 questionnaire respondents who provided the phone numbers were contacted. Only 10 of them replied but three students are not available; hence, there were 7 participants for the interviews.

Two participants were interviewed face to face; the conversations were noted down and recorded. The remaining were interviewed via the Internet with the help of Facebook video call, no recording tape was provided in the application so the research made attempt to take as much note as possible.

Step 4: Collecting, quantifying and qualify data

Quantitative data

Quantitative data was collected and analyzed by Anova to answer to research question 3, the descriptive statistics and frequency were applied to answer question 1 and 2.

Qualitative data

Data was reviewed and coded. Each participant was coded as *GPA_number* (i.e. GPA_1)

3.6. Data analysis

Step 1: Processing quantitative and qualitative data

Firstly, the researcher codes data. For the questionnaire, data were coded in SPSS program. Regarding the interview, data were put in different themes.

Secondly, the internal reliability of the questionnaire was checked by using Cronbach's alpha estimation in SPSS program. However, the unordered categorical data such as the school year or reasons to take the GPA were not examined by this software. As the purposes and type of questions in each part were different, the internal reliability testing of each part was processed separately. The result is shown as below.

Part 1: Students' perspectives toward the GPA

Table 3.4. Reliability Statistics part 1

Reliability Statistics		
Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.654	.659	6

Part 2: Students' general English activities (this part filters the cases of non-GPA test-takers)

Table 3.5. Reliability Statistics part 2

Reliability Statistics		
Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.633	.644	16

The open-ended question “How many hours do you spend doing those activities?” was excluded from the test, nearly a quarter of respondents do not give answer and the answers vary wildly.

Part 3: Students' specific test preparation (this part filters the cases of non-GPA test-takers)

Table 3.6. Reliability Statistics part 3

Reliability Statistics		
Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.704	.588	15

At all parts, the Cronbach's Alpha is larger than 0.6, which indicates that the questionnaire is acceptable (Stephanie, 2014). The statistics shows that the questionnaire was sufficiently reliable for analysis.

Step 2: Interpreting quantitative and qualitative data

Items are categorized following related questions

- Research question 1

- Research question 2 and question 3

To answer question 1 and 2, various types of questions were employed in the questionnaire including multiple choice questions (MCQs), rating scales and open-ended questions. The frequency of each option was calculated and presented in percentage terms and/ or numerical terms. Data were then demonstrated by means of tables, pie charts and bar chart.

The research question 3 is to identify the differences between the group of novice students and the seniors; hence, one-way ANOVA test was applied. The extent of the difference is decided by ANOVA Sig. of which the value is $0 < \text{Sig.} < 1$. If ANOVA Sig. < 0.05 or Sig. = 0.05, there is a difference and vice versa. Phạm (2016), in his article, presented in great details how to find ANOVA Sig. and how to interpret it.

Step 3: Finalizing the report

The data which were computed and illustrated in the previous steps were interpreted. By analyzing data, students' perspective towards the test and their reflection toward test preparation were discovered.. Besides, the explanations for those views might be revealed from interviews with students. Additionally, by comparing the test preparation activities of the respondents, the differences between two groups were identified.

Chapter summary

This chapter described the methodology applied in this study with the specific context at the beginning. A mixed method using both survey and interview was employed in order to get both the overall trend and deep understandings of the issues. The sample was first stratified, then the questionnaires were delivered to each stratum following clustering samples. There were 123 responses in which 7 respondents took part in the interview round. The procedure of collecting, processing and analyzing data were also thoroughly presented.

CHAPTER 4: RESULTS AND DISCUSSION

4.1. Students' attitude toward the graduation proficiency assessment (GPA)

This section presents the results on students' perspectives toward the GPA and their test anxiety based on the questionnaire. Additionally, further findings from interview data are discussed. As described in the methodology chapter, research question 1 "What are students' attitude toward the graduation proficiency assessment?" can be answered by six-questionnaire items:

- Why do you take the VSTEP instead of IELTS, TOEFL, TOEIC or other compatible tests as the graduation test? (Q3)
- Do you think the graduation proficiency assessment (GPA) will *motivate* students to improve their English? (Q6)
- Do you think the GPA will *raise* the proficiency of English of university students? (Q7)
- Do you think GPA will *measure* students' English proficiency *accurately*? (Q8)
- Do you think the GPA test result will be *useful* in helping you find a job? (Q9)
- How do you feel about having to take the GPA? (Q5)

The answers for these questions are presented below.

4.1.1. Students' decision to take the GPA

Respondents stated their reasons to choose GPA by answering an open-ended question. This type of question helps the scholar to collect numerous

opinions, but its disadvantage is the low response rate; in return, there were only 57 answers which are categorized into 7 groups as in figure 4.1 below.

Regarding the nature of the test, many respondents (31.2%) believed that the GPA was the easiest proficiency test in comparison with TOFEL, IELTS, and TOEIC. Besides, the GPA was perceived to be convenient for registration and be at reasonable price for non-English majored students at the School of Law.

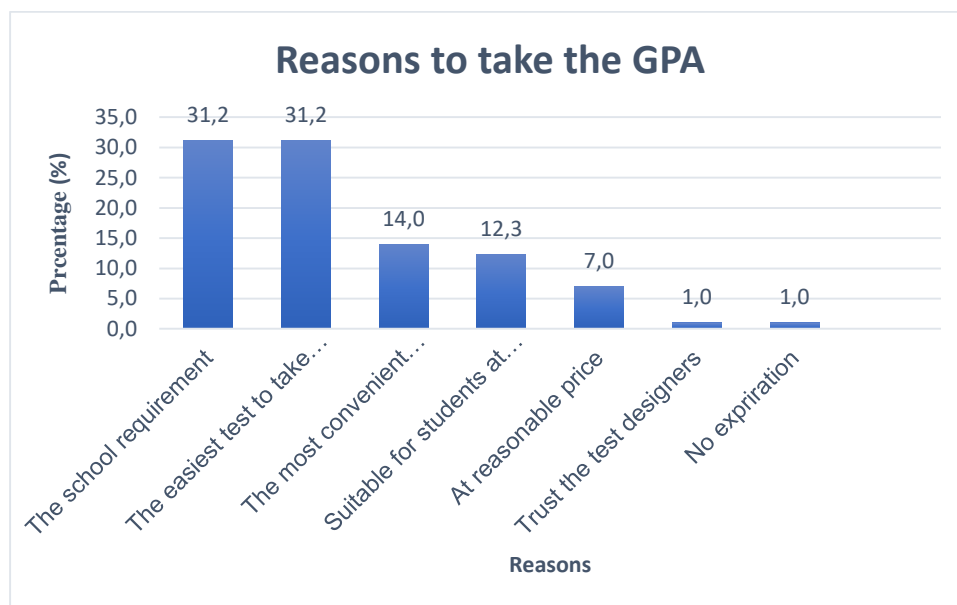


Figure 4.1. Reasons to take the GPA

Nearly a third of respondents (31.2%) stated that they took the test as the school's requirement. It was revealed in the interview that many students did not know they could take other tests to transfer their scores.

For personal reasons, one student stated that she trusted the test designers in the University of Languages and International Studies (ULIS).

In summary, the GPA seems to be less unchallenging but more suitable for students compared with the alternative proficiency tests. Respondents took test for it is the school's precondition and the GPA is considered as the easiest choice.

4.1.2. Whether the test would motivate students to improve their English

Respondents were asked to answer yes/ neutral/ no question to the question "Do you think the graduation proficiency assessment (GPA) will

motivate students to improve their English?” (Q6). The finding shows that more than a half of respondents (56.3%) consented to the motivating function/ impact of the GPA, whereas just more than a fifth (22.7%) showed the opposite opinions; the rest accounting for 21% had no opinion. This result indicates a positive trend among the students of the Law School that the GPA would motivate students to learn English. The subsequent interview gave some explanations for this thinking.

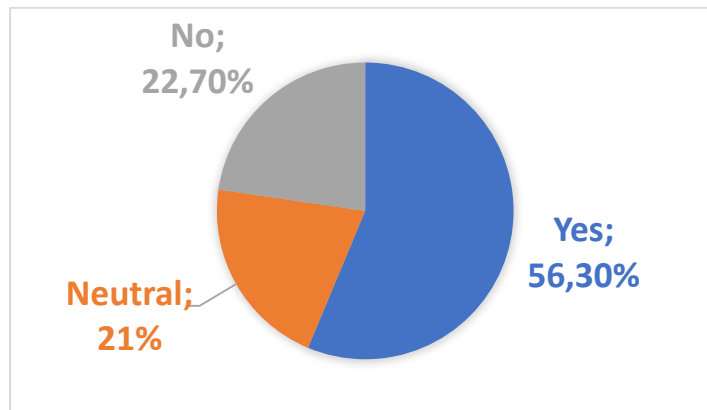


Figure 4.2. Students’ perception toward test motivation

Students thought passing the test was crucial to get their degrees and graduate from the university or to cover the scores in English courses; hence, they were motivated to study for the GPAS.

“Like other tests, the GPA forces me to study however boring English is to me”
(GPA_1)

“if it had not been for the GPA, I would not have studied English” (GPA_4)

One participant indicated that the GPA stimulated her to study the language further.

*“The test informs me about my English proficiency, so I know where I am, what are my strengths and weaknesses. Therefore, I know what’s more I should learn.”*GPA_2

Regarding the claim that the test does not provoke learning, some participants found their wants, the test’s level, and test importance as the culprit.

“I didn’t want to learn English, instead I wanted to spend time learning Chinese. The test just stressed me out but not inspire me” GPA_4

“I find the GPA not too challenging; besides, I am confident with my English level, so I don’t spend much time to learn it” GPA_3

In short, most respondents found the test motivate them to improve their English; whereas, some students considered the test unimportant and unchallenge to stimulate learning.

4.1.3. Whether the test would raise students’ English proficiency

Respondents answered yes/ neutral/ no to the question “Do you think the VSTEP will raise the proficiency of English of university students?” The table below displays the distribution of students’ perspectives toward the raising proficiency function of the test. Overall, the result shows enthusiastic responses. 63.1% participants answered positively while only 17.5% answered “no”.

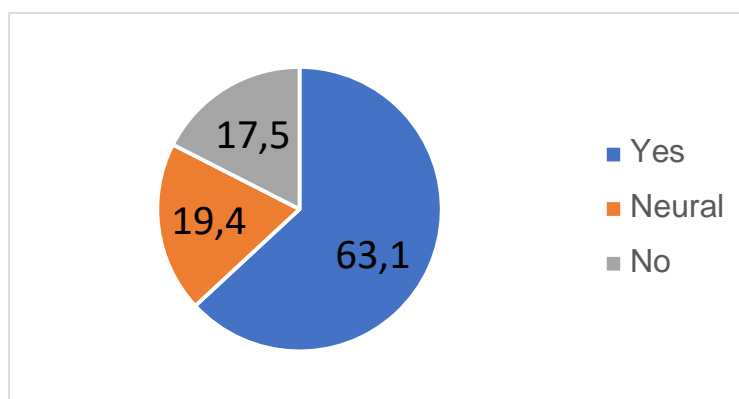


Figure 4.3. Students’ perception toward test function of raising English proficiency

It is quite surprising that most interviewees did not agree with the findings. Many respondents including the low and higher English competence ones claimed that their English level did not improve at all owing to the test.

“I find the test easy so I merely spend time for it. I think my English proficiency does not raise at all.” GPA_3

“Getting B1 is quite hard for me so I go to an English center in which I am provided with help by teachers who are familiar with the test format. I do not learn the basic knowledge such as grammar, vocabulary, and skills but I am given the templates to speak and write. However, they are unpractical as I cannot speak or write for communication purposes. Therefore, I do not think the GPA raise my competence at all.” GPA_4

Even though many students insisted on gaining no more knowledge from learning for the test, throughout the interview, the researcher could spot their minor improvements.

“In the speaking test, I understood most phrases and ideas that the teachers said, but I could not speak fluently. When I got out of the room, numerous ideas flashed through my mind. I regretted not telling them.” GPA_5

This student found herself know nothing in English when first entering the university. If she could speak a little English even it was not natural, it should be considered as one of her improvements.

In short, while the questionnaire data demonstrates a vast number of the respondents perceive the test as a means to raise students English proficiency, interview participants showed pessimistic views toward the issue. The cause for this contrast might be that the interviewees did not see their progress in language learning.

4.1.4. Whether the test would measure students’ proficiency accurately

Respondents answered yes/ neutral/ no to the question “Do you think VSTEP will *measure* students’ English proficiency *accurately*?” to indicate their attitude toward the test accuracy. It should be noted that these respondents are not testing or language experts to have an insight view; hence their answers could not be regarded as the evidence to prove the test is accurate or not. Instead, the question purpose is to figure out how students perceive the test accuracy. The term “accuracy” refers to the quality of the test, such as face validity and reliability; since these students are not supposed to have knowledge of the jargon, an alternative term was acquired.

The chart below shows the proportion of students’ responses toward test accuracy. Different opinions dispersed among students; while nearly a half had no option or left the answer blank; the patterns for “yes” and “no” were almost equal (30,1% versus 27%).

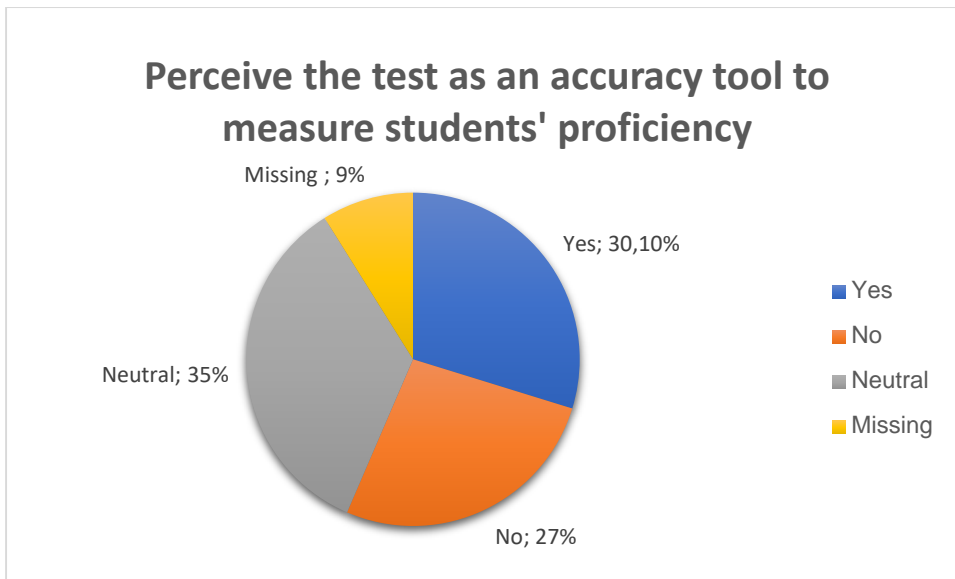


Figure 4.4. Students' perception of the GPA's accuracy

The interview data conveyed an insight explanation for these answers. Many students did not rate the test accurate as they were not familiar with the test.

"I am not clear about the GPA format and structure so I cannot give you an answer"

Some thought that this question should be answered by teachers and expert instead of students. As they found themselves not entitled and trustworthy to deem the test.

"I don't know, I am not an expert."

For students who held positive perspective, the test structure and content seemed to persuade them.

"Well. I think the GPA is fairly accurate as it assesses four skills listening, speaking, reading and writing like many other proficiency tests."

"I think the test would be more suitable for Vietnamese people. Other proficiency tests namely IELTS, TOEFL, and TOEIC cause difficulties as they are influenced by foreign mindset. The GPA, on the other hand, sets more familiar contexts." GPA_1

However, regarding the reliability of individual skills, many students expressed personal doubts toward MCQs and assessing speaking skills.

“I think reading, listening, and writing tests are quite okay, but I am not sure about speaking skills, the scores vary following the examiner.” GPA_2

“I don’t think multiple choice questions can assess students’ levels accurately, as they can choose randomly.”

Others believed that the overall score did not reflect the test-taker English proficiency.

“I know many people at my university who got scores equivalent to B1 or even B2 level but cannot speak or write in English as the description of those levels. For example, if she gets a B1 certificate, she should be able to talk about common topics or have simple conversations, right? My friend just can greet like hi, or goodbye.” GPA_3

“I don’t think my score reflects my current level at all. You know, I got an extremely low score in listening and considerably high score in reading; overall, I passed the GPA. I know my final score is the average score of four skills, but I feel like I am not really qualified. My listening score is too low to B2 level. I think the test designers should give a restriction to each skill. For example, all my skills have to above 3 or 4, I don’t know, to pass the B2 level.” GPA_4

To summarize, the non-English major students did not evaluate whether the GPA was accurate or not but gave their opinions toward the test. Although the distribution seems to equal among “yes” and “no”, the students’ perspectives on the test’s downside should be taken into serious consideration.

4.1.5. Whether the test’s results would be useful in supporting job application

Respondents chose yes/ no/ neutral to question “Do you think the VSTEP test result will be useful in helping you find a job?” to indicate their opinions toward the test usefulness. Consistent with the previous section, a large number of respondents (34,9%) did not have answers or stayed neutral. The option “yes” made up 39.9% while 25.2% respondents chose “no”. In other words, optimistic

respondents outnumbered the opposite side 1.5 times. On that account, there is a slight tendency that respondents perceived the test usefulness positively.

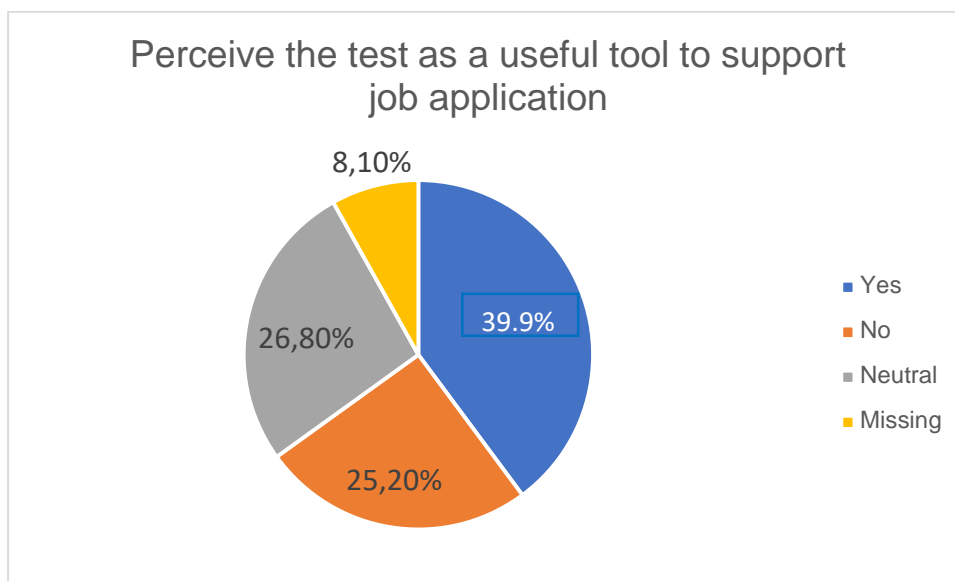


Figure 4.5. Students' perception toward the test usefulness

These following interview excerpts would give explanations why the GPA is useful or not. Firstly, the GPA was considered advantageous in job application on the account of the great needs of English competent workers in Vietnamese labor market.

“English gives me more chance when seeking for jobs. There are a certain percentage of recruiters who are interested in the GPA. I just don’t know exactly how many of them.” GPA_6

On the other hand, many believed the unpopularity of the GPA was the cause for its uselessness.

- *“The GPA at my school which is now called VSTEP is not really popular with the employers. If it were well-known, I believe the test would be more useful.”* GPA_6

- *“Even many of my friend in VNUH who are required to take the GPA don’t know its existence; consequently, the test also sounds unfamiliar to the recruiters. I think they gives more attention to international proficiency test such*

as IELTS, TOEFL or TOEIC rather than the GPA. Hence, I don't consider the GPA useful." GPA_1

The interviewee also pointed out that many employers might be skeptical about the test result.

"I cannot use the GPA result to apply for a job, as I don't think the employers would trust the test score."

In summary, respondents tend to find the GPA beneficial for job application. Although the test might be helpful for recognizing respondents' English proficiency, not many recruiters would concern and accept this test. The respondents suggested that the GPA should be more prevail and more accurate to support students in finding jobs.

4.1.6. Test anxiety

Students answered question (5) "How do you feel about having to take the GPA?" by ranking from 1 to 5 from extremely worried to extremely confident. The figure 4.6 below displays the percentage of respondents in each state of feeling. Overall, the most chosen state of mind was "neutral" which meant not too anxious nor too confident (36.6%) while the least popular feeling was "extremely confident" which made up only 2.4%. As it can be seen from the chart, the number of anxious students outnumbered the self-assured ones; to be more specific, it was 48% versus 12.5%. The statistics also indicates that students tend to feel rather nervous of taking the GPA with mean = 2.39 and SD = 1.

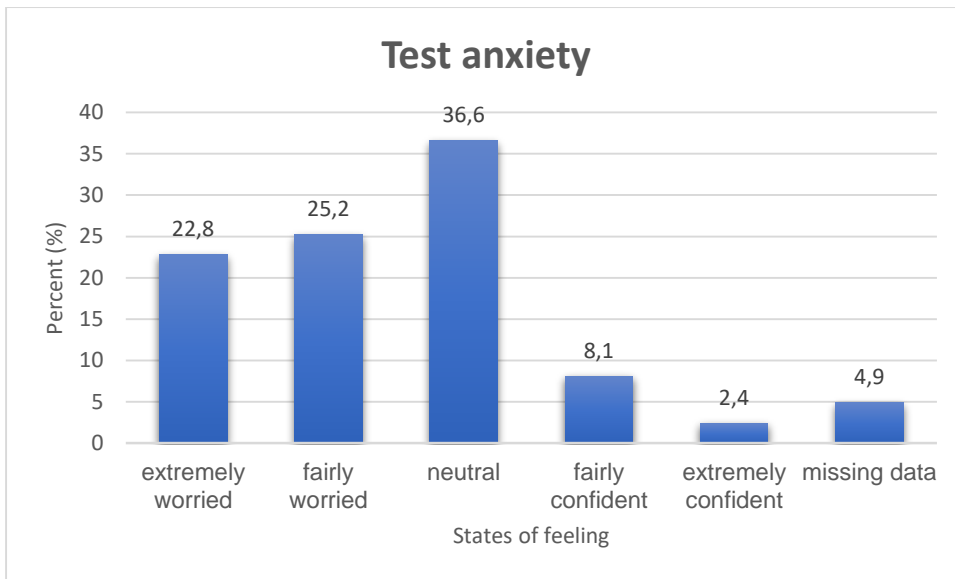


Figure 4.6. Respondents' test anxiety

The interview data showed that most respondents were not confident as they were not competent at English. The reasons for respondents' anxiety came from their fear of wasting time and money, not being able to prepare thoroughly and testing pressure itself.

"I paid a fortune for the English center and for the GPA fee, if I failed I would have to take the course and study again. Such a waste of time and money. Therefore, I am under pressure of passing the test." GPA_5

"I was worried about not knowing what was exactly in the real test. I felt like going into a battle unprepared." GPA_3

"For me, any tests or exams cause me nervous, waiting for the test to come is a stressful period." GPA_1

However, some interviewees admitted that they were not desperately anxious as the test was not very challenging.

In short, respondents are likely to be slightly nervous before the test due to the anxiety of losing cash and time, the uncertainty in test structure, and the nature of testing itself. Besides, respondents were not too confident or too worried as they were aware of their ability compared with the GPA.

4.2. GPA washback on students test preparation for the graduation proficiency assessment

Since the study aims to examine students who have taken the GPA, those who sit for other tests were excluded. There were only 102 questionnaires remaining. In studying the washback on students test preparation, The dimension of general/ specific following the paper of Cheng et al. (2004) and (Stoneman, 2005) was applied

4.2.1. The general English studies

A multi-response question and an open-ended one were employed to investigate (1) types of activities, and (2) the time spent on these activities of respondents. These questions help to identify the general English activities of the students.

The multi-response questions asked respondents to choose the activities that they apply to study English such as watching English TV programs, listening to English radio/ song, reading English books/ newspaper/ magazines, writing diaries in English and Speaking English with teacher. This list of activities was adapted from Stoneman's (2005) questionnaire. In 102 responses, two respondents stated to play games in order to study English, this information is not shown in the chart.

As it can be seen from the chart, the three most common activities are related to receptive skills such as listening to English radio/ songs (79.4%), watching TV programs (63,7%) and learning English on the Internet (54.9%). Productive skills activities including speaking and writing are less commonly applied (<20%). The extend that each student take up general activities is presented by how many ways or methods that each employs. The table below show the number of English learning activities taken up by the respondents ranking from one to eleven activities

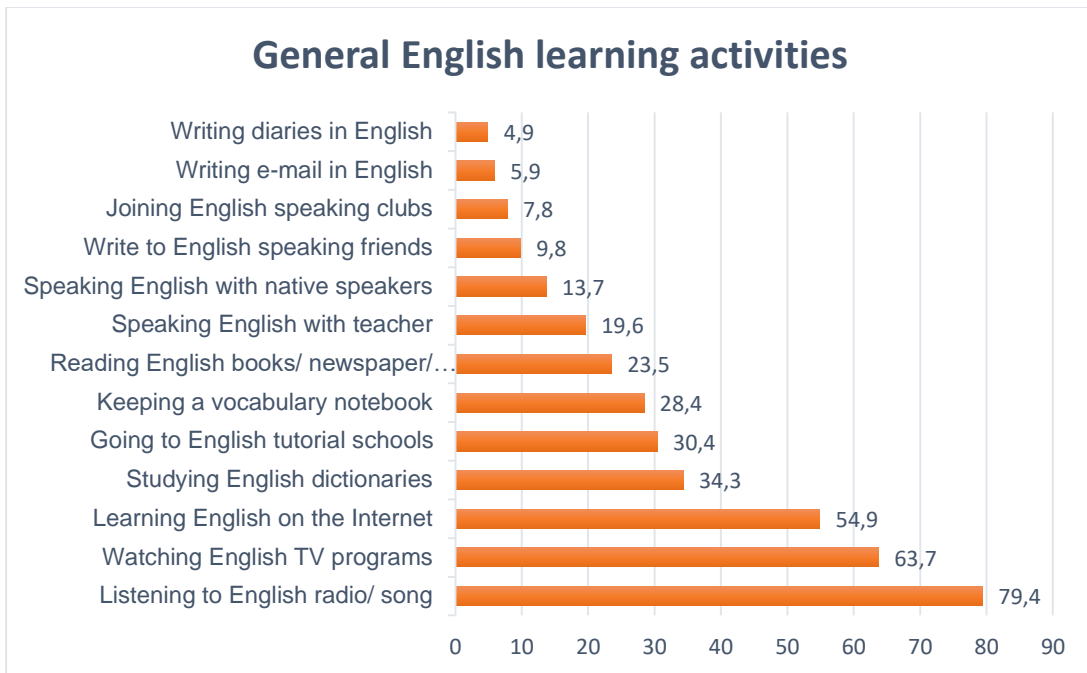


Figure 4.7. Students general English learning activities

While the number of activities undertaken widely varied among the respondents, in average they engaged in three to four activities (mean = 3.76). Likewise, the percentiles show that 50% of the respondents participated in four activities. Four out of thirteen is trivial number what infers the respondents did not join in various learning activities

Table 4.8. Number of General learning activities

Mean	3.7647
Mode	2.00
Std. Deviation	2.30014
Minimum	.00
Maximum	11.00
Percentiles 25	2.0000
Percentiles 50	4.0000
Percentiles 75	5.0000

Table 4.9 below shows how many hours each week that students spend time learning English. It should be noted that this table is not highly reliable but just for reference since many students complaint that they could not remember or count the exact how many hours they study English a week. As their major is

not English, they do not learn it continuously and constantly. As a result, many blanks were left and many respondents may have made up a number to satisfy the researcher. In 102 responses, solely 76 students stated the numbers of hours they devoted to learning which ranges enormously from 0 to 40 hours. Although in most cases, students claimed to study 2 hours weekly, the average number was much higher at 5.9 hours. This figure is later compared with the statistic of students' test specific preparation in the next part.

Table 4.9. Numbers of hours per week for General English activities

N	Valid	76
	Missing	26
	Mean	5.90
	Median	4.00
	Mode	2
	Std. Deviation	6.832
	Minimum	0
	Maximum	40
	Sum	449

This finding goes in line with the study of Stoneman on the Hong Kong' students out of class activities. (Stonean, 2006). Additionally, very few students claimed to continue those activities when preparing for the test.

4.2.2. Test-specific English activities

The researcher asked 4 questionnaire items from question 12 to 15 to investigate how students prepare for the test. (Appendix 2). Question 12 enquired students about their preparation status particularly whether they have taken any steps or not. The next three questions asked about the time, learning strategies and the content when students study for the test either in the past, at the present or in the future as they take the GPA at different times. Question 13 gave the respondents three options regarding the number of hours they spend studying the test per week: 0 -2 hours, 2 – 5 hours, and more than 5 hours. Students learning strategies such as seek help from teacher, access the GPA's website, study with friends or self-study were investigated in question 14. Question 15 helped the researcher identify to what extent the respondents studied the test-like materials.

4.2.2.1. Test preparation status

Regarding test preparation status, the chart 4.10. below compare if students have prepared for the test between two groups. Overall, the third and last year students had a clearer view of their provision with 54.8% “yes” answer compared with 11.3% “yes” answer of the counterpart. A third of respondents in the first and second year students group did not do any preparation while the figure was only more than 40% for the other group. The finding is consistent with the hypothesis that the first and second year students tend to have less preparation than the third and fourth year students. The ANOVA test also approved the difference (Welch Sig. = 0.00) between the two group: the former group are unlikely to have taken the studying (mean = 1.87), while the latter tend to have made several preparations (mean = 1.44).

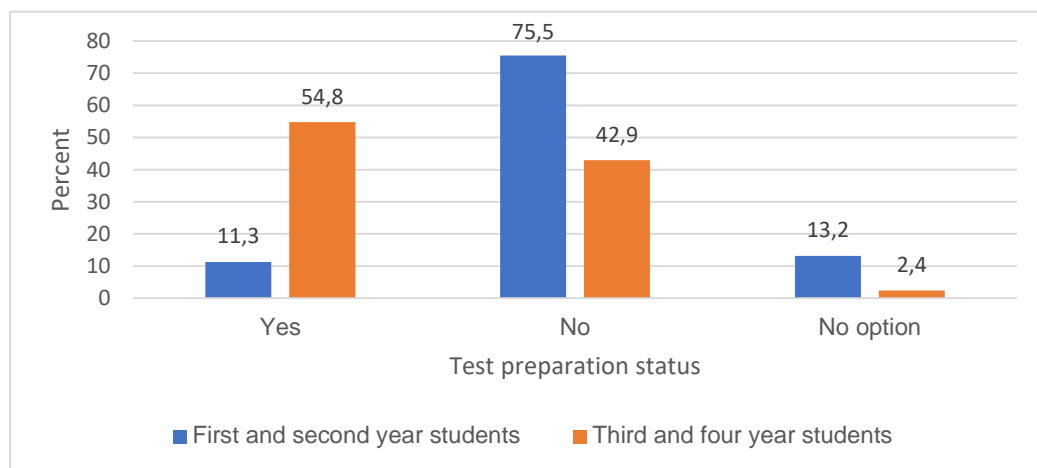


Figure 4.10. Students test preparation status

4.2.2.2. Time allotment

Regarding to the amount of time that students prepare for the test, figure 4.11. below shows a wide range of periods among the respondents.

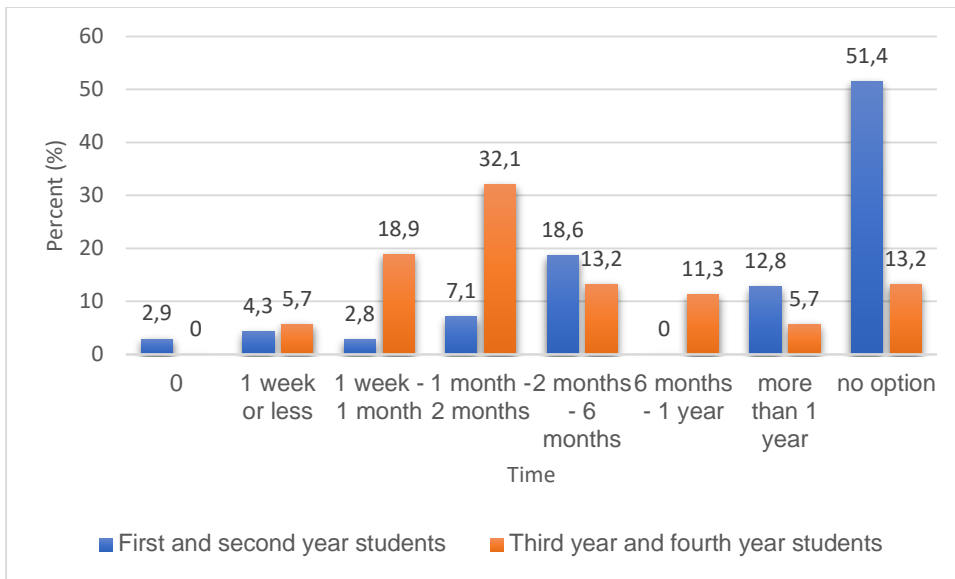


Figure 4.11. Time to prepare before the test

The first-year students and sophomore had a high rate of uncertainty toward the preparation period, as up to 51.4% respondents had no clue when they would prepare for the test. In contrast, the figure was just under 15% “no option” answer for the third and last year students. While the former group had a various test preparation periods ranging from no time (2.9%), a week or less (4.35%), a week to one month (2.85%), one month to two months (7.1%) to more than a year, the figure for the latter group was more concentrated around one month to six months (64.2%). To illustrate, the largest proportion of respondents studied for the test at the period of “one month to two months” with 32.1%, followed by “one week to one month” phase with 18.9%; only 13.2% respondents claimed to study up to six months. Overall, the senior group are likely to spend a short period to study for the test, whereas the novice ones are uncertain. This finding agrees with Shih (2007)’s study on washback on students learning. He also finds that students spend few days or weeks or up to two months but there is little long-term effort.

Table 4.12. below shows in the period of their GPA preparation, how many hours would the respondents study on weekly basis. A nearly equal number of respondents studied one to two hours and two to five hours per week, 28.4% and 27.5% respectively, but only 17.6% respondents spent more than five hours a week. Mean = 1.85 showing that in average respondents spent two to four hours

for learning for the test. Besides, the ANOVA test showed that there was no significant difference between two groups.

Table 4.12. Time for test preparation of respondents per week

How many hours per week will/do/did you prepare for the test?				
	Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1-2 hours (1)	29	28.4	38.7
	2 - 5 hours (2)	28	27.5	37.3
	more than 5 hours (3)	18	17.6	24.0
	Total	75	73.5	100.0
Missing	9	27	26.5	
Total	102	100.0		

Looking back at question 11, it was estimated that in average respondents spent up to 5.9 hours per week for English learning activities. This was slightly higher than test preparation time (two to four hours). In the questionnaire data, 16 respondents took up to 10, 20 or even 40 hours to study English but only spent less than five hours for test preparation. The interview data shed some light on in this issue. Some interviewees considered that the test itself could not help them improve language proficiency; thus, apart from time for test preparation, they also learn English by watching films, reading newspaper or listening to music.

However, test provision activities were prioritized when the GPA was closer as those activities might help the respondents obtain better scores.

“Well, I think if people study the test in the long term, they would do some out of class activities like those. However, when the test is closer, most people focus on practice for the test only.” GPA_7

In short, more seniors have taken steps to prepare for the GPA than the novice ones who seem to be unsure about test provision. However, there is “little long-term effort” as period of two months or less is the most favored among the third and four-year students. With regards to time allotment, time for specific test preparation activities are negligibly lower than that for general activities. Students had different ideas for test preparation. Some thought preparing for the test was “a small piece in the whole picture of language learning”; thus, they may continue out-of-class activities. Others argued that less or even no general

English activities could be used to study for the test. However, they all concurred that when the test was closer, more test specific activities would be employed.

4.2.3. Test preparation strategies

Question 14 contains four sub-questions which ask whether respondents (a) seek for help from English teacher (b) access the GPA website (c) study in group with friends (d) self-study to find out what students' learning strategies are. The number of learning strategies are counted and presented in the table 4.2.1_4 below. The choices range from 0 methods to maximum 4 methods while 2 is the most common selection. On average, the participants take up about 1.7 learning strategies (mean = 1.696; SD = 1.21).

Table 4.13. Number of learning strategies taken by respondents

N	Valid	102
	Missing	0
Mean		1.6961
Median		2.0000
Mode		2.00
Std. Deviation		1.20879
Minimum		.00
Maximum		4.00
Sum		173.00
Percentiles	25	1.0000
	50	2.0000
	75	3.0000

Regarding the test preparation strategies, the figure 4.2.2_4 below illustrates the percentage of different methods taken by two group of students. There is no remarkable distinction between two group according to ANOVA test (Sig. > 0.05).

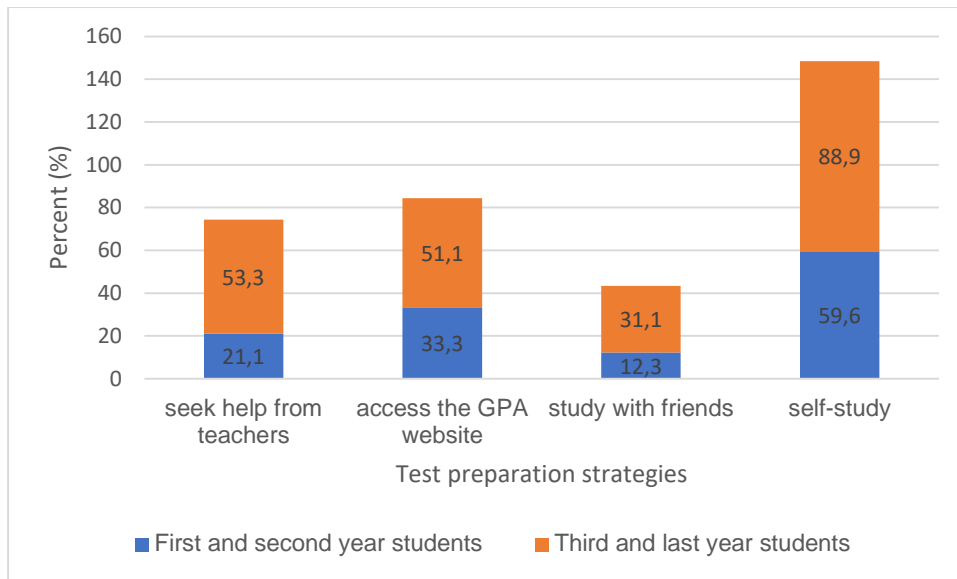


Figure 4.14. Test preparation strategies

As it can be seen from the bar chart, self-study is the most applied method with 88.9% first and second year students and 59.6% the third and fourth year. Following that, the proportions of respondents seek help from teachers and access the GPA website are roughly equal.

Only few would study with friends to prepare for the test.

These following interview excerpts may of help for further understanding of each learning strategies. Those who prepare for the test themselves seem to be active in findings materials and engaged in studying but encounter the problems of what to learn.

- *“I don’t study at any center, but self-study at home. I find the information of test date, test format and sample tests. Then I look for the exercises which are equivalent to the real test to do. However, I could not prepare much because I am not sure whether those exercises are compatible to the test. I think the school should provide more free sample tests.”* GPA_1

- *“I studied for the test by myself, as I think I was competent enough. It would be best to have teachers to instruct you what to do but I didn’t have money. Finding an English partner at this school seems to be impossible as they are not English-major, not confident, and comfortable to speak. Well, it was hard for me to find the right materials, but I had no other choice”* GPA_4

Some go to English teacher for help and get more test-like materials.

- *“I went for teachers who helped students with the GPA, and only completed the materials that teachers provided. The teacher gave us the exercises as the test format. For speaking and writing, we were given templates, we just needed to combine and add some details.”* GPA_5

- *“I just study at the class listening to the teachers, doing exercises, taking notes and so forth but I do not learn anything at home”* GPA_4

Accessing the GPA website for test information such as time, location, test format and past paper test seems to be common; however, little materials and examples are found. Consequently, test-takers are not sure about the test format.

- *“I found some exercises but I wasn’t not assure they were similar as the GPA. It caused me certain anxiety.”* GPA_4

Studying with friends is the least chosen strategy as many students consider themselves not competent; besides, English is not their major; thus, students have little chance to work in group, but invest more in their major group work.

In summary, two groups of students seem to be similar in choosing learning strategies. While self-study is the most common method, studying with friends is not favored. Many people also go for English centers and GPA website.

4.2.2.4. Learning contents

Question 15 including 8 statements required the respondents to rate the amount of time they study the test-like materials from never (1) to a lot of time (4). The close-to-test content contains the past paper test and similar exercises as test format such as reading and listening for MCQs, speaking about familiar topics, choosing one in three option, writing a letter, and writing an essay. Figure 4.15 below shows the amount of time respondents devote to each test preparation activity.

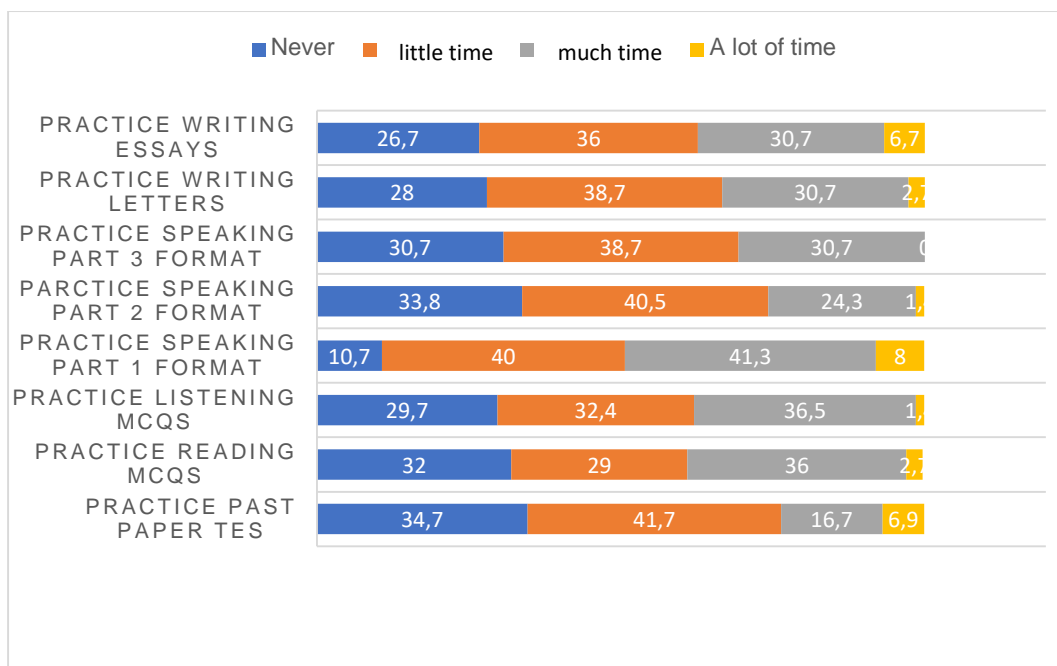


Figure 4.15. Time for each test preparation activity

Overall, while “never” practicing the test format was the least common choice, most of the respondents devoted “little time” and “much time” for the mentioned activities. Besides, means ranging from 1.97 – 2.1 and SD equaling around 0.8 indicate that in general respondents spend little time doing the test-like exercises.

Explaining for the shortage of test-like practice, many respondents stated that they were unfamiliar with test format, the materials for GPA were hard to find if they did not study at English centers or had other supports.

“My friend told me to read and listen for MCQs in the books FCE, practice writing letters and essays following the IELTS Trainer. Without her, I don’t know what to study”. GPA_3

Other reason was that the respondents also spent time on learning general English namely grammar and vocabulary

“Even though the test doesn’t include grammar and vocabulary tasks, I still feel an urge to improve these areas so as to get a higher score.”

Regarding the difference between two group, the ANOVA test only showed different means in “practice reading MCQs and practice speaking by choosing one option of three”. However, the stacked bar 4.16 below compares test preparation activities between first year & second year students (FYSY) and

third year & last year students (TYLY) showing the dissimilar tendency in most activities.

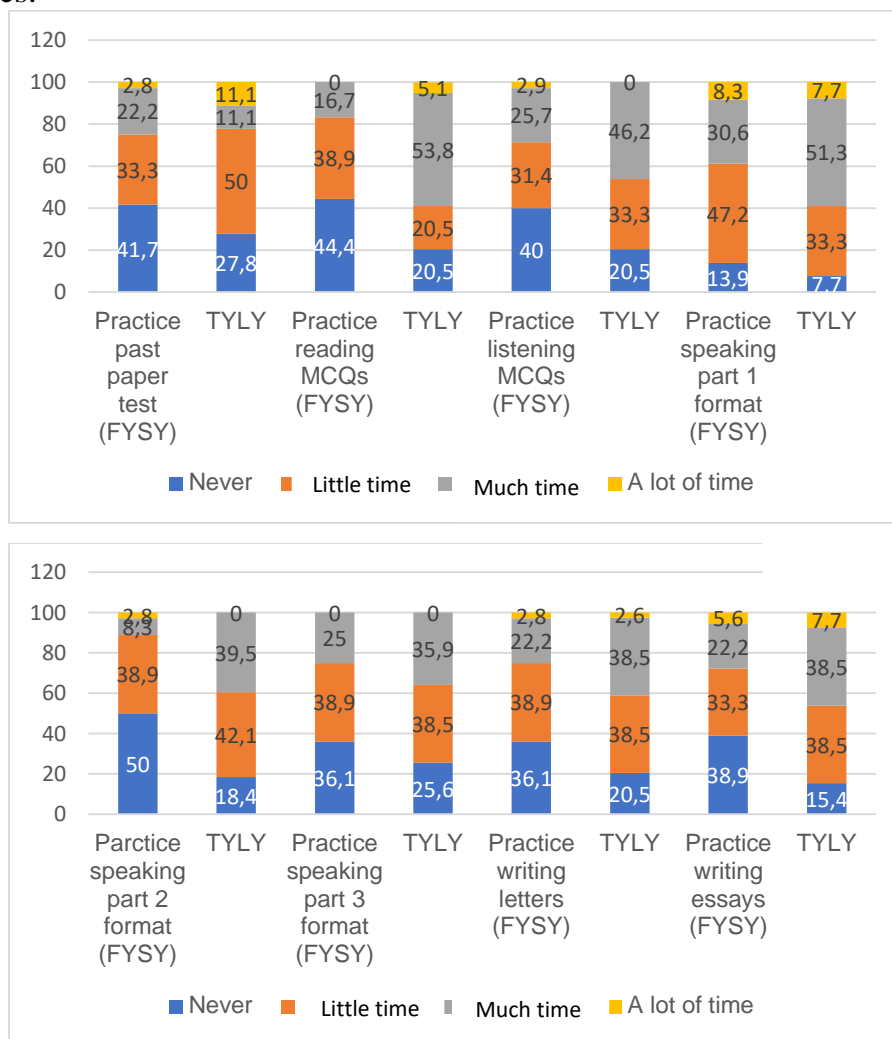


Figure 4.16. Difference between two groups

For the mentioned activities, there is an apparent trend that the first year and second year students who “never” practiced the test approximately outnumbered the counterpart. For instance, in practice speaking part 2, up to a half of FYSY respondents never did it, the figure was only 18.4% for the TYLY students. As a result, the time that the seniors did the test-like materials was over that of the FYSY. To be more specific, while more FYSY students spent “little time” for the test, the TYLY had a tendency to practice the test “much time”.

These findings indicate that the GPA has exert specific washback on the test-takers; the closer the test is, the more specific test preparation activities. However, the washback is just minimal since respondents do not make long-term effort to prepare for the test; various TYLY students study for the test in two

months or less while most FYSY respondents even have no plan for the provision. In addition, that students do not devote “a lot of time” for preparation infers that the test do not induce a high degree of washback. This is consistent with the study of Shih (2007), in which students prepare for the test but with little effort; thus, he concludes that the General English Proficiency Test (GEPT) has limited washback on students.

To conclude, students spend “little time” on the materials which are similar to the test. Perhaps, they have limited understanding of the test format or they are likely to learn the basic knowledge. There is different inclination between two groups of students that the seniors make more effort to prepare for the test. Consequently, the GPA has little washback on students’ test preparation.

Chapter summary

This chapter has reported and interpreted quantitative and qualitative data to answer three research questions. Research question one is answered by analyzing 6 attitudinal questionnaire items. The findings show that in general, students have positive thinkings toward the test. For research question 2 and 3, respondents who has no intention to take the GPA is excluded from the study as the aim of these questions are identify students’ test preparation activities and the difference between the freshmen and seniors; six behavioral questions are studied. Quantitative data are presented in tables, pie charts and bar charts for analysis; then the findings are further described by interpreting qualitative data.

PART V: CONCLUSION

The previous chapter has discussed the findings and interpretation from quantitative and qualitative data to answer three research questions. This conclusion chapter is going to summarize the main findings, evaluate the study and give some recommendations for the students and teachers at the researched school. Besides, the limitation of the study and some suggestions for further studies are also put forward.

5.1. Summary of the findings and discussion

5.1.1. Students attitude toward the test

Most students at the School of Law take the GPA before graduating instead of registering other proficiency tests for several reasons. Firstly, the test is considered undemanding, convenient and economical to take. Besides, some respondents have little knowledge of the school policy and reckon the test as a compulsory one.

In general, most respondents have positive thinking toward the GPA. Regarding the test as a tool to motivate and raise students' English proficiency, the majority of the respondents show optimistic responses. Most students feel encouraged to study English to pass the test so that they could meet the school's requirement, cover their English scores, and recognize their English levels. In terms of test usefulness, there is a slight tendency toward the positive as numerous respondents reckon that English plays a critical role in job finding process and the GPA could show those recognitions. With regards to attitude to test accuracy, no trend could be drawn from the survey.

Although the GPA seems to get credit for the mentioned reasons, **some respondents hold negative views toward the test.** For few interviewees, the test goes against their wants, does not change high proficiency students as well as motivate low competent ones. Likewise, there are several claims that either the easiness or the practicing to the test do not raise students' English levels. The test reliability and accuracy are also questioned for marking speaking may differ from different examiners, MCQs might cause the random choices, and some find

their friends unqualified but still pass the exam. As the unpopularity of the test among the recruiters and the seem to low of value certificate, a couple of respondents do not regard the test useful for their job application.

For test anxiety, a multitude of respondents are nervous before the test because of the unpopular format and test-taking pressure. However, most of them do not experience the extreme worry as the GPA is consider not too demanding.

5.1.2. Washback on students' test preparation

With regard to the test preparation behaviors, the general English learning activities are studied. In average, **the respondents engage in a modest number of activities**, only three to four out of thirteen ones in the survey. Time spent for those activities vary from 0 to 40 hours per week. However, those behaviors are claimed to lessen when the test is closer. Regarding the test specific preparation behaviors, there is **a larger number of the third year and last year respondents taking steps to prepare for the GPA than the first year and second year students**. While the former has made effort for test preparation in approximate 2 months, most of the students in the latter group are uncertain about the study plan. Despite those differences, **the two groups appear to have the same the test preparation strategies**. A vast number of students study for the test by themselves for saving money purpose. In the second place is seeking for help from English teachers and accessing the GPA website which provide learners with more knowledge of the test. While English center offers students with excessive test materials, the GPA website is perceived to provide poor content for provision.

In terms of test-like content for preparation, **few respondents practice the test-like materials and they are likely to spend little time studying instead of a great quantity of time**. There is a clear difference between two groups of students shown by the report. While up to a half of the FYSY respondents have never practice as the test format, the figure is much smaller in the TYLY group. As a result, the latter group study more test similar tasks mentioned. Additionally, while more FYSY students practice the test only “little time”, the counterpart spends more time for test-like preparation. In short, the

seniors have more preparations and do more activities close to the test, while the freshmen are likely to make little attempt toward the GPA. It can be inferred that the GPA has little washback on students' preparation.

5.2. Conclusion and implication

This study aims at investigating the washback back of GPA on students' test preparation and their attitude toward the test. To achieve this aim, the researcher has initially reviewed the literature of washback and the proficiency test to obtain both theoretical and empirical background. The frameworks of Cheng et al. (2004) and (Shih, 2007) are then adopted. Following that, the methodology has presented in details the context, participants and research search design including how the data are collected, analyzed and interpreted. Both survey and interview are employed to answer research questions. The findings and discussion about students' attitude toward the test and washback on students' test preparation are then presented in chapter 4.

5.3. Recommendations

The understanding of washback on students' test preparation and students' attitude toward the test is of great help for students, teachers, and test administration. From the findings, the researcher proposes some minor suggestions as following:

Firstly, students should build up their language proficiency gradually so that they would not end up cramping for exam and consequently studying just test-like materials.

For teachers, besides test-similar materials and templates, they should also provide students with more practical exercises, so after taking, the students could apply the knowledge in daily situations.

Last but not least, test administrators should take the feedbacks of the students into serious consideration. The complaints such as the GPA is insufficiently popular to support the application process, the test seems to be unreliable and invalid to students and the materials for the GPA is particularly rare should be fully resolved.

5.4. Limitations of the study

Although the study has been carefully carried out, limitations are unavoidable. Firstly, owing to time limit, the study was conducted on a small population size, only the non-English major at the School of Law. Therefore, to generalize results for a larger group, there should be more participants. Secondly, this study collected second hand data, opinions of the students, which may not reflect the actual performance. Hence, for further study, the observations or experiences method could be employed. Finally, some other aspects and dimensions of washback have not been concerned for the limitation of the research.

5.5. Suggestions for further research

Due to the limitations of this research, there are rooms for further studies. The following are some suggestions:

- Washback on students on larger scale with multiple methods
- Washback on teachers and others participants
- Washback on all level of Vietnamese educational system

REFERENCES

- Bachman, L. F. (2000). Modern language testing at the turn of the century: assuring that what we count counts. *Language Testing*, 17(1), 1–42.
- Bailey, K. M. (1996). Working for washback: a review of the washback concept in language testing. *Language Testing*, 13(3), 257–278.
<https://doi.org/10.1177/026553229601300303>
- Barnes, M. (2016). The Washback Effects of the TOEFL iBT in Vietnam. *Australian Journal of T Teacher Education*, 41(7).
- Brown, H. D., & Abeywickrama, P. (2010). *LANGUAGE ASSESSMENT Principles and Classroom Practices* (2nd ed.). NY: Pearson Education.
- Caine, N. A. (2005). EFL examination washback in Japan: investigating the effects of oral assessment on teaching and learning. The University of Manchester.
- Carr, N. T., Nguyen, T. N. Q., Nguyen, T. M. H., Nguyen, T. Q. Y., Nguyen, H. L. T., & Nguyen, T. P. T. (2016). Systematic Support for a Communicative Standardized Proficiency Test in Vietnam.
- Cheng, L., Watanabe, Y., & Curtis, A. (2004). *WASHBACK IN LANGUAGE TESTING Research Contexts and Methods*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Damankesh, M., & Babaii, E. (2015). The washback effect of Iranian high school final examinations on students' test-taking and test-preparation strategies. *Studies in Educational Evaluation*, 45, 62–69.
- Frederiksen, J. R., & Collins, A. (1989). A Systems Approach to Educational Testing. *American Educational Research Association*, 18(9), 27–32.

- Green, A. (2007). Washback to learning outcomes: a comparative study of IELTS preparation and university pre-sessional language courses. *Assessment in Education: Principles, Policy & Practice*, 14(1), 75–97. <https://doi.org/10.1080/09695940701272880>
- Hamp-Lyons, L. (1997). Washback, impact and validity: ethical concerns, *14*, 295–303.
- Hughes, A. (1989). *Testing for L*
anguage Teacher. Cambridge: Cambridge University Press.
- Linn, R. L. (1983). Testing and instruction: links and distinctions.
- M.Elshawa, N. R., Heng, C. S., Abdullah, A. N., & Rashid, S. M. (2016). Teachers' Assessment Literacy and Washback Effect of Assessment. *International Journal of Applied Linguistic & English Literature*, 135-141.
- Nguyen, H. (2014). Đề án thiết kế định dạng đề thi năng lực tiếng Anh từ bậc 3 đến bậc 5 cho đối tượng sau trung học phổ thông (Vietnamese Standardized Test of English Proficiency). Đại học Ngoại Ngữ - ĐHQGHN.
- Nguyen, P. N. (1997). *Washback effects of the international English language testing system at the Vietnam National University, Hanoi*. University of Melbourne, Australia.
- Nguyen, T. L. (2016). *Impacts of Vietnamese Standardized Test of English Proficiency (VSTEP) on the first year students' English language learning at University of Languages and International Studies (ULIS), Vietnam Nation University, Hanoi (VNU)*. Hanoi: University of Languages and International Studies.
- Pan, Y.-C. (2014). Learner Washback Variability in Standardized Exit Tests. *Teaching English as a Second or Foreign Language*, 18(2).

- Pan, Y.-C., & Newfields, T. (2012). Tertiary EFL proficiency graduation requirements in Taiwan: A study of washback on learning. *Electronic Journal of Foreign Language Teaching*, 9, 108–122.
- Pham, T. H. (2010). *Evaluation of washback effects of ESP end - term test for second-year students at UEB-VNUH on teachers at Faculty of English - ULIS - VNUH*. Hanoi: H. : University of Languages and International Studies.
- Phạm, L. (2016). Phân tích khác biệt trung bình One-Way ANOVA trong SPSS.
- Popham, W. J. (1987). The Merits of Measurement-Driven Instruction. *Phi Delta Kappan*, 68, 67–82.
- PrimeMinister. (2008). *Decision 1400/QĐ-TTg*. Hanoi
- Shih, C.-M. (2007). A New Washback Model of Students' Learning. *The Canadian Modern Language Review*, 64(1), 135–162.
- Shohamy, E., Donitsa-Schmidt, S., & Ferman, I. (1996). Test impact revisited: washback effect over time. *Tel Aviv University*, 13(3), 298–317.
- Skehan, P. (2008). Language testing. Part 1. *Language Teaching*, 21(4), 211–221.
- Stephanie. (2014, August). Cronbach's Alpha: Simple Definition, Use and Interpretation. *Statistics How To*.
- Stoneman, B. W. H. (2005). The impact of an exit English test on Hong Kong undergraduates: a study investigating the effects of test status on students' test preparation behaviors. Pao Yue-kong Library PolyU Hong Kong.

- Sukyadi, D., & Mardiani, R. (2011). The Washback Effect of the English National Examination (ENE) on English Teachers' Classroom Teaching and Students' Learning. *k@ta*, 96-111.
- Tsagari, D. (2011). Washback of a high-stakes English exam on teachers' perceptions and practices, *19*, 432–445.

APPENDIX

Appendix 1

WASHBACK OF THE PROFICIENCY ASSESSMENT AS EXIT REQUIREMENT ON NON-ENGLISH MAJOR STUDENTS AT A LAW SCHOOL

FELTE, ULIS, VNU

Questionnaire

I am Huong from QH2014.F1.E3 of ULIS, VNU, Hanoi and I am carrying a research on *“Washback of the proficiency assessment as exit requirement on non-English major students at a law school”* which gives a closer look on how law students prepare for the graduation proficiency assessment (GPA) and their attitudes influenced by the test. Your response to this questionnaire is of great help for the final product, which might assist test-designers develop a fairer and more valid test. I hope you provide trustworthy answers as it is only way to ensure the validity of the study. Your information will be kept confidential under any circumstances. Thank you very much for your contribution!

There is a following interview to discuss the exit test further. The participants will get 50.000 if they finish both the questionnaire and the interview. In case, you would like to attend the interview, please leave your email and phone number below. The researcher will contact you as soon as possible.

I appreciate your help in completing both the questionnaire and the interview, thank you once again.

Your sincerely,

Bui Thanh Huong

PART 1: BACKGROUND INFORMATION

Your name:

Phone number:Email:

1. Did you/ do you intend to take the graduation proficiency assessment (GPA) offered by your university?

Yes No

2. You are: First year student Third year student
 Second year student Fourth year student

PART 2: PERSPECTIVES TOWARD THE GRADUATION PROFFICIENCY TEST (GPA)

3. Why do you take the VSTEP instead of IELTS, TOEFL, TOEIC or other compatible tests as the graduation test?

.....

4. How do you rate your overall English language ability? (*Please tick the box in the scale below*)

very poor _____ *excellent*

1 2 3 4 5

5. How do you feel about having to take the VSTEP? (*Please tick the box in the scale below*)

extremely _____ *extremely*

worried 1 2 3 4 5 *confident*

6. Do you think the graduation proficiency assessment (GPA) will motivate students to improve their English?

No Yes No option

7. Do you think the VSTEP will raise the proficiency of English of university students? (please tick one box)

No Yes No option

8. Do you think VSTEP will measure students' English proficiency accurately? (please tick one box)

No

Yes

No option

9. Do you think the VSTEP test result will be useful in helping you find a job?

(please tick one box)

No

Yes

No option

PART 3: GENERAL ENGLISH ACTIVITIES

10. Which of the following out-of-class activities do you do? (Please tick the box/ boxes that apply) 0. is an example

<input checked="" type="checkbox"/>	0. Watching English TV programs	<input type="checkbox"/>	7. Speaking English with teachers
<input type="checkbox"/>	1. Watching English TV programs	<input type="checkbox"/>	8. Speaking English with native speakers
<input type="checkbox"/>	2. Listening to English radio/music	<input type="checkbox"/>	9. Going to English tutorial schools
<input type="checkbox"/>	3. Reading English books/ newspaper/ magazines	<input type="checkbox"/>	10. Joining English speaking clubs
<input type="checkbox"/>	4. Writing diaries in English	<input type="checkbox"/>	11. Learning English on the Internet
<input type="checkbox"/>	5. Write to English speaking friends	<input type="checkbox"/>	12. Studying English dictionaries
<input type="checkbox"/>	6. Writing e-mail in English	<input type="checkbox"/>	13. Keeping a vocabulary notebook

Others.....

11. Approximately how many hours per week do you normally spending on doing those activities outside class? _____ hours

PART 4: TEST PREPARATION

12. Have you done anything to help you to prepare for the test? (Please tick the box)

NO

YES _ Please specify the activity or activities

.....

13. a. How long before the GPA did/will/do you start preparing for the test?

(give a number and circle year/ month/ week/ day) i.e. 2 weeks

..... year/ month/ week/ day.

b. How many **hours per week** did/will/do you study for the GPA?

1 - 2 hours 2 - 5 hours > 5 hours

14. What are your learning strategies to prepare for the GPA? (*Please tick V the column YES/ NO*)

Learning strategies	NO	YES
a. Seek help from English teachers		
b. Access the GPA website		
c. Study with friends		
d. Self-study		

15. Do you do these activities when preparing for the VSTEP? (*Please tick the box/ boxes that apply*) 0. is an example

0. I practice the past paper test	✓
1. I practice the past paper test	
2. I practice items similar in the test format	
a. I practice reading MCQs questions If yes, where _____	
b. I practice listening MCQs questions If yes, where _____	
c. I practice speaking familiar topic about myself and my surrounding (study, work, family, friends, hobbies, sports, fashion, etc.)	
d. I practice speaking by choosing an option in three and giving reasons	
e. I practice speaking by developing ideas for a particular topic (e.g. benefits of reading, etc.)	
e. I practice writing letter	
d. I practice writing essays (discussion, opinion, problems and solutions, advantages and disadvantages essays)	
I study vocabulary for common topics in the test	

-----This is the end-----

Appendix 2

TÁC ĐỘNG CỦA BÀI ĐÁNH GIÁ NĂNG LỰC TIẾNG ANH CHUẨN ĐẦU RA (CĐR) ĐỐI VỚI SINH VIÊN KHÔNG CHUYÊN KHOA LUẬT PHIẾU THĂM DÒ Ý KIẾN

Các bạn thân mến,

Tôi là Hường, sinh viên khóa QH2014.F1.E3 trường Đại học Ngoại Ngữ - ĐHQGHN. Tôi đang tiến hành một nghiên cứu về “**Ảnh hưởng của bài đánh giá năng lực Tiếng Anh chuẩn đầu ra đối với sinh viên không chuyên**”. Đề tài đưa ra cái nhìn cụ thể hơn về sự chuẩn bị của sinh viên khoa Luật cho kỳ thi Đánh giá Năng lực tiếng Anh chuẩn đầu ra (CĐR) và nhận thức của sinh viên về kỳ thi. Câu trả lời của bạn đóng góp một phần không nhỏ cho sự hoàn thiện nghiên cứu này, có thể giúp đỡ những người thiết kế bài thi phát triển bộ đề công bằng và chính xác hơn. Tôi mong bạn có thể cung cấp câu trả lời chân thật nhất vì đó là cách duy nhất để đảm bảo tính chính xác cho nghiên cứu này. Thông tin của bạn sẽ được giữ bí mật trong mọi trường hợp. Cảm ơn sự đóng góp của bạn!

Sau phiếu thăm dò ý kiến, sẽ có một **buổi phỏng vấn** sâu hơn về bài thi chuẩn đầu ra. Người tham gia sẽ được **nhận phần quà tương đương 50.000 vnd** nếu hoàn thành cả phần khảo sát và phỏng vấn. Trong trường hợp bạn muốn tham gia buổi phỏng vấn, hãy để lại địa chỉ email và số điện thoại bên dưới. Tôi sẽ liên lạc để xếp lịch với bạn sớm nhất có thể.

Tôi rất trân trọng sự giúp đỡ của bạn trong quá trình khảo sát và phỏng vấn.

Xin cảm ơn bạn một lần nữa.

Thân ái,

Bùi Thanh Hường

PHẦN 1: THÔNG TIN CĂN BẢN

Tên bạn là (không bắt buộc).....

Số điện thoại: Email:

1. Bạn có định thi/ đã thi CĐR ở trường cho yêu cầu tốt nghiệp không?

Có Không

2. Bạn là:

Sinh viên năm nhất Sinh viên năm ba

Sinh viên năm hai Sinh viên năm tư

PHẦN 2: THÁI ĐỘ CỦA SINH VIÊN ĐỐI VỚI BÀI THI CHUẨN ĐẦU RA

3. Tại sao bạn chọn thi CĐR thay cho thi IELTS, TOEFL, TOEIC hay bài thi khác cho yêu cầu tốt nghiệp? (Nếu bạn chọn bài thi khác, hãy nêu rõ đó là bài thi gì, tại sao bạn lại chọn)

.....

4. Bạn tự đánh giá khả năng Tiếng Anh của mình như thế nào? (Hãy đánh dấu V vào ô trống bên dưới)

rất tệ _____ tuyệt vời

1 2 3 4 5

5.

Bạn thấy thế nào khi phải thi CĐR? (hãy tích vào một ô trống dưới đây)

Rất lo lắng _____ Rất tự tin

1 2 3 4 5

6. Bạn có nghĩ Kỳ thi Đánh giá Năng lực Ngoại ngữ (CĐR) sẽ tao động lực để sinh viên cải thiện Tiếng Anh? (hãy tích vào ô trống)

Không Có Không ý kiến

7. Bạn có nghĩ là kỳ thi CĐR giúp làm tăng chuẩn Tiếng Anh của sinh viên đại học? (hãy tích vào ô trống)

Không Có Không ý kiến

8 Bạn có nghĩ bài thi CĐR đánh giá trình độ tiếng Anh của sinh viên một cách chính xác? (hãy tích vào ô trống)

Không

Có

Không ý kiến

9. Bạn có nghĩ là kết quả bài thi CĐR có ích trong việc giúp đỡ bạn tìm được một công việc tốt hơn? (hãy tích vào ô trống)

Không

Có

Không ý kiến

PHẦN 3: HOẠT ĐỘNG HỌC TIẾNG ANH CHUNG CHUNG

10. Hoạt động ngoài giờ nào mà bạn thực hiện để học Tiếng Anh? (Hãy đánh dấu vào ô trống phù hợp với bạn)

0. là một ví dụ

<input checked="" type="checkbox"/>	0. Watching English TV programs (<i>Xem chương trình TV</i>)	<input type="checkbox"/>	7. Speaking English with teachers (<i>Nói Tiếng Anh với giáo viên</i>)
<input type="checkbox"/>	1. Watching English TV programs (<i>Xem chương trình TV</i>)	<input type="checkbox"/>	8. Speaking English with native speakers (<i>Nói Tiếng Anh với người bản xứ</i>)
<input type="checkbox"/>	2. Listening to English radio/music (<i>Nghe radio/ nhạc Tiếng Anh</i>)	<input type="checkbox"/>	9. Going to English tutorial schools (<i>Đến trung tâm Tiếng Anh</i>)
<input type="checkbox"/>	3. Reading English books/ newspaper/ magazines (<i>Đọc sách/ báo/ tạp chí Tiếng Anh</i>)	<input type="checkbox"/>	10. Joining English speaking clubs (<i>Tham gia câu lạc bộ Tiếng Anh</i>)
<input type="checkbox"/>	4. Writing diaries in English (<i>Viết nhật ký Tiếng Anh</i>)	<input type="checkbox"/>	11. Learning English on the Internet (<i>Học Tiếng Anh trên mạng</i>)
<input type="checkbox"/>	5. Write to English speaking friends (<i>Viết/ chat với bạn người bản xứ</i>)	<input type="checkbox"/>	12. Studying English dictionaries (<i>Học từ điển Tiếng Anh</i>)
<input type="checkbox"/>	6. Writing e-mail in English (<i>Viết email bằng Tiếng Anh</i>)	<input type="checkbox"/>	13. Keeping a vocabulary notebook (<i>Giữ một quyển sổ từ vựng</i>)

Khác:

11. Một tuần bạn dành bao nhiêu thời gian để học Tiếng Anh?

..... giờ.

PHẦN 4: HOẠT ĐỘNG HỌC TIẾNG ANH CỤ THỂ ĐỂ CHUẨN BỊ CHO CĐR

12. Bạn đã làm gì để chuẩn bị cho kỳ thi CĐR chưa? *(hãy tích vào ô trống)*

CHƯA

RỒI

13. a. Bạn dành bao nhiêu thời gian để chính thức chuẩn bị cho CĐR?

(hãy đưa ra một con số và khoanh vào năm/ tháng/ tuần/ ngày)

..... **năm/ tháng/ tuần/ ngày.** (Ví dụ:
2 tuần)

b. Bạn dành **bao nhiêu tiếng một tuần** để chuẩn bị cho CĐR?

1 - 2 tiếng 2 - 5 tiếng > 5 tiếng

14. Chiến lược học bạn áp dụng để chuẩn bị cho CĐR như thế nào? *Đánh dấu tích (V) vào cột CHƯA/ RỒI*

Các chiến lược chuẩn bị thi	CHƯA	RỒI
a. Tìm sự giúp đỡ từ giáo viên Tiếng Anh (đến trung tâm tiếng Anh/ gia sư/...)		
b. Truy cập trang web của CĐR để tìm hiểu đề thi như cấu trúc, đề mẫu, v.v		
c. Học nhóm tiếng Anh với bạn		
d. Tự học		

15. Bạn học những mảng gì để chuẩn bị cho CĐR? *Đánh giá thời gian bạn thực hiện các hoạt động dưới đây bằng việc chọn mức độ từ 1 (không bao giờ) đến 4 (rất nhiều thời gian)*

1 – không bao giờ

2 – một ít thời gian

3 – nhiều thời gian

4 – rất nhiều thời gian

1. Luyện tập đề đã thi.	1	2	3	4
2. Luyện tập câu hỏi giống trong định dạng đề thi.	1	2	3	4
a. Đọc và trả lời câu hỏi trắc nghiệm. Tài liệu nào (nếu có).....	1	2	3	4
b. Nghe và trả lời câu hỏi trắc nghiệm. Tài liệu nào (nếu có).....	1	2	3	4
c. Luyện nói các chủ đề thân thuộc: về bản thân và về môi trường xung quang (học tập, công việc, gia đình, bạn bè, sở thích, thể thao, v.v.)	1	2	3	4
d. Luyện nói bằng cách chọn một trong ba lựa chọn rồi giải thích lý do	1	2	3	4
e. Luyện nói bằng cách phát triển các ý cho một chủ đề (vd lợi ích của việc đọc, v.v)	1	2	3	4
f. Luyện tập viết thư	1	2	3	4
g. Luyện tập viết luận (thảo luận, vấn đề và giải quyết, lợi ích và tác hại)	1	2	3	4

Kết thúc phiếu khảo sát

Cảm ơn bạn rất nhiều!

Nếu bạn muốn tham gia buổi phỏng vấn, nhớ điền **số điện thoại** và **email**. Bạn sẽ được nhận **phần quà tương đương 50.000 vnd** khi buổi phỏng vấn kết thúc. Nếu có câu hỏi, liên lạc với tôi qua số điện thoại (01669602811)

Appendix 3

The interview questions – English version

I. BACKGROUND

1. Do/ did you take the VSTEP? Why?
2. When did you/ will you take the test?

II. ATTITUDE TOWARD THE TEST

2. At that time, did you consider VSTEP as an important test?/ Do you consider VSTEP as an important test?

- No (why not)
- Yes (important in what way? And how important on a 1 (+) to 5 (+++) scale)

3. At that time, who did you think were interested in the VSTEP test scores/ test results?

Why were they interested/ not interested?

- Yourself
- Parents
- University
- Employer

4. Did/ do you feel motivated by the test to put more effort into improving your English?

- No (why not)
- Yes (how motivating on a 1 (+) to 5 (+++) scale)

5. Did the GPA psychologically influence you during your preparation for the GPA? If yes, how, were/ are you worried, or were/ are you confident? (1 to 5 scale, 1 = extremely worried, 5 = extremely confident) Why?

6. Did/ do you think VSTEP assess students' English levels accurately? Why?

III. LEARNING PREPARATION

a. Timing

7. How long did/ will you spend to prepare for the test in total? Per week?
8. Do you think it was/ is enough? Why?

b. Content/ Method

9. Tell me through the process and stages of how you prepared/ prepare for the VSTEP:

When did you did what?

Who you went to?

What resources you used?

10. What would you otherwise do if you did not have to pass the GEPT as one of the degree requirements? Would you take the GEPT? Would your test preparation be different?

Appendix 4

Các câu hỏi phỏng vấn

I. BỐI CẢNH

1. Bạn có tham gia thi VSTEP không? Tại sao?
2. Bạn đã làm bài kiểm tra khi nào?

II. THÁI ĐỘ ĐỐI VỚI KỲ THI

2. Vào thời điểm đó, bạn có coi VSTEP quan trọng không?
 - Không (tại sao không)
 - Có (quan trọng theo cách nào? Và tầm quan trọng của thang điểm 1 (+) đến 5 (+++))
3. Vào thời điểm đó, bạn nghĩ ai quan tâm đến điểm thi / kết quả thi VSTEP? Tại sao họ quan tâm / không quan tâm?
 - Bản thân bạn
 - Cha mẹ
 - Trường đại học
 - Nhà tuyển dụng
4. Bạn có nghĩ bài kiểm tra thúc đẩy sinh viên nỗ lực hơn trong việc cải thiện tiếng Anh không?
 - Không (tại sao không)
 - Có (cách thúc đẩy tỷ lệ 1 (+) đến 5 (+++))
5. GPA có ảnh hưởng tâm lý đến bạn trong quá trình chuẩn bị không? không? Nếu có, làm thế nào, là / bạn đang lo lắng, hoặc là / bạn có tự tin? (Tỷ lệ 1 đến 5, 1 = cực kỳ lo lắng, 5 = cực kỳ tự tin) Tại sao?
6. Bạn có nghĩ rằng VSTEP đánh giá chính xác trình độ tiếng Anh của học sinh không? Tại sao?

III. CHUẨN BỊ HỌC

a. Thời gian

7. Bạn đã dành bao nhiêu thời gian để chuẩn bị cho bài thi? Mỗi tuần?
8. Bạn có nghĩ rằng đó là / là đủ? Tại sao?

b. Nội dung / Phương pháp

9. Cho tôi biết qua quá trình và các giai đoạn về cách bạn chuẩn bị / chuẩn bị cho VSTEP:

Khi nào bạn đã làm gì?

Bạn đã đến ai?

Bạn sử dụng tài nguyên nào?

10. Bạn sẽ làm gì nếu bạn không phải vượt qua GEPT như một trong những yêu cầu mức độ? Bạn có tham gia GEPT không? Chuẩn bị kiểm tra của bạn có khác không?

Appendix 5

VSTEP Format

VSTEP Format	Listening	Reading	Writing	Speaking
Time	40 minutes	60 minutes	60 minutes	12 minutes
Number of questions	3 sections 35 MCQs	4 passages 40 MCQs	2 writing tasks	3 parts
Task	Test-takers listening to short conversation, instruction, announcement, dialog and discussion; then answer MCQs	Test-takers reads for passages about different topics, with total approximately 1900-2500 words; then answer MCQs after each text	Task 1: Write a 120 – word letter. Task 2: Write 250 – word essay on a given topic	Part 1: Social interaction Test-taker answers 3-6 questions in two different topics. Part 2: Solution discussion Test-taker receives a situation with 3 solutions. He/she has to give opinion on the best solution and criticize other two. Part 3: Topic development Test-taker speaks about a

				given topic, using given ideas or his/ her own ideas. Part 3 ends with some discussion questions.
Purpose	Assess different listening subskills from level 3 to 5: listen for details, listen for main ideas, listen for opinion and purpose of speakers and infer information.	Assess different reading subskills from level 3 to 5: reading for details, reading for main ideas, reading for opinions, attitudes of the writers, infer information and guess the word meaning from the text,	Assess letter writing skills and persuasive writing skills	Assess different speaking skills: Social interaction, discussion and topic development

e