

THE ROLE OF LEARNERS' TEST PERCEPTION IN CHANGING ENGLISH LEARNING PRACTICES: A CASE OF A HIGH-STAKES ENGLISH TEST AT VIETNAM NATIONAL UNIVERSITY, HANOI

Nguyen Thuy Lan^{*1}, Nguyen Thuy Nga²

*1. Academic Affairs Department,
VNU University of Languages and International Studies,
Pham Van Dong, Cau Giay, Hanoi, Vietnam*

*2. VNU University of Education,
144 Xuan Thuy, Cau Giay, Hanoi, Vietnam*

Received 10 October 2019 Revised 15 November 2019; Accepted 20 December 2019

Abstract: Among various factors influencing foreign language learning, learners' perception of a high-stakes language test plays a crucial part, especially when the test serves as a threshold for their university graduation. In this study, the researcher tested a washback effect model by focusing on test-takers' perception of the high-stakes test VSTEP in terms of test familiarity, test difficulty and test importance. On a sample of 751 Vietnamese learners of English at Vietnam National University, structural equation model was employed to validate the conceptual model. The analytical methods of Exploratory Factor Analysis (EFA), Confirmatory Factor Analysis (CFA) and Structural Equation Modeling (SEM) were used for analysis. Our empirical findings revealed that VSTEP seems to have had a pervasive impact on the participating students. Senior students' evaluations of VSTEP acted as the largest factor in constituting the participants' perception of VSTEP. There are positive links between test pressure and test familiarity with students' goal setting and study planning as well as their selection of learning content and materials. Meanwhile, the pressure from the test had no effect on students' seeking opportunities to practice with foreigners, and test familiarity did not influence students' choice of study methods and exam preparing strategies. The emerging patterns from the data also suggested that participating students preferred test-oriented learning content and activities at the cost of interactive English practices for real-life purposes.**

Key words: learners' perception, high-stakes tests, washback effect, test-oriented, SEM

1. Introduction

The academic regulations of Vietnam National University, Hanoi (VNU) attached

to Decision No. 5115/QĐ-ĐHQGHN on December 25th, 2014 clearly states that non-English-major students are required to submit evidence of English proficiency level 3 or B1 (CEFR - Common European Framework for Reference) for graduation. Launched by VNU University of Languages and International Studies in 2017, Vietnamese Standardized Test of English Proficiency 3 (VSTEP 3) is

* Corresponding Author. Tel.: 84-928003530

Email: lanthuy.nguyen@gmail.com

** This research is funded by VNU University of Education (UED) under the project number QS.18.09.

a standardized test designed to measure the English proficiency of VNU undergraduate students and to determine whether their English-language ability meets the requirements of level 3 or B1 as a graduation condition.

In accordance with the university curriculum, students are eligible to take VSTEP 3 only after they have completed three English modules (General English 1, 2 and 3). VSTEP 3 is held twice a year: in June, at the end of the spring semester, and in December, at the end of the fall semester. Like most of the CEFR-based tests, VSTEP consists of four sections: listening, reading, writing and speaking.

While students and teachers are under high pressure of achieving the learning outcomes upon graduation, and a new standardized test is used as an official instrument to measure students' language proficiency, the question is whether the test has made changes to students' English learning practices.

In the past several decades, the impact of tests has been the subject of considerable attention from educators and researchers — especially in the field of language testing worldwide. However, there is a dearth of empirical evidence in regard to test effects in Vietnamese language education context. In this article, we initially aimed to explore and analyze some effects of students' perception of the VSTEP 3 as a high-stakes test on their English learning practices.

2. Literature review

2.1. High-stakes tests

According to Minarechova (2012), a high-stakes test is no longer a new educational phenomenon. It has become an integral part of the educational system in many countries.

Madaus (1988) defines a high-stakes test as a test whose results are used to make important decisions affecting the students, teachers, managers, the school and the community in its geographical area. The purpose of a high-stakes test is to link learner's results in standardized tests with the outcome requirement for the completion of an educational level; and in some cases, it is the base to review the wage increase, or sign the long-term work contract with teachers (Orfield & Wald, 2000).

In line with the aforementioned definitions, Vietnamese Standardized Test of English Proficiency 3 – VSTEP 3 is a high-stakes test as it is used as the official language proficiency tool to make an important decision: whether students can graduate from their university and be prepared for job seeking.

2.2. Washback effects

Research in the field of testing and assessment asserted that tests, especially high-stakes tests, had great impacts on teaching and learning activities. These effects are commonly considered “washback effects”. This concept has been defined in various ways in the history of research. Alderson & Wall (1993) defines “washback effects” (washback or backwash) as the effect of the test back into the teaching and learning process. This concept derives from the view that the testing and assessment can and should orient the teaching and learning process. According to Alderson and Wall (1993), washback effects only refer to the behaviors of learners and teachers within the classroom when influenced by a particular test. To clarify the degree and extent of the test, many authors have distinguished between the washback effect and the impact of the test. Wall (1997) states that “the effect of the test “is” ... any effect of the test on the individual, the policy in the classroom, the school, the educational

system or the whole society”; meanwhile, the washback effect of the test only refers to the “effects of the test on teaching and learning” (p. 291). Similarly, Shohamy (2001) suggests that the effect of washback effect is a component of test impact. The impact of the test takes place on a social or an educational institution, but the washback effects influence learners and teachers. The washback effect is also considered an aspect of the value of a test and is referred to as “consequential validity”, which emphasizes the “consequence” of examinations, testing and assessment on previous teaching and learning (Messick, 1996).

2.3. Related studies on the washback of language tests and learners' test perception on English learning

Hughes's (1993) model is a pioneer washback model which discusses the complex process of washback occurring in actual teaching and learning environments. Hughes (1993) distinguishes between participants, processes and products in both teaching and learning, recognising that all three may be affected by the nature of a test. The participants, including students, teachers, administrators, materials developers, and publishers are those whose perceptions and attitudes toward their work may be affected by a test. The process is any action taken by the participants that contributes to the learning process. The products refer to what is learned and the quality of the educational outcomes. According to Hughes (1993), a test will first influence the participants' perceptions and attitudes, then how they perform, and finally the learning outcomes.

Kirkland (1971) stated that students are the primary stakeholders in testing situations as it is the student “whose status in school and society is determined by test scores and the one

whose self-image, motivation, and aspirations are influenced” (p. 307). In the same line, Rea-Dickins (1997) recognized students' significant role in the process of test washback; he also added that “their views are among the most difficult to make sense of and to use” (p. 306). In the literature of washback effects, researchers, however, have tended to focus on test impact on teaching activities, whereas studies on students have met with scant attention. Furthermore, in rare student-related research, most studies have focused on academic factors, whereas students' affective conditions have been neglected. It is, therefore, important to directly assess how students feel about the test and how their perception of the test affects their English learning.

Etten, Freebern & Pressley (1997) conducted an interview-based study with an aim to detail college students' beliefs about the examinations they face. The researchers interviewed those closest to the exam preparation process, those who make the decisions about when, how, and what to study, college students themselves. The conclusions that emerged from several rounds of questioning were a complex set of beliefs about the examination preparation process. According to Etten, Freebern & Pressley (1997), there were a number of external factors that influence test preparation, and the most significant could be named as instructors, exam preparation courses, social environmental variables, physical environment, test-related materials, all of which could undermine or facilitate studying.

In his extensive literature review, Kirkland (1971) concluded that tests could have impacts on a range of factors related to students, including self-concept, motivation, level of aspiration, study practices, and anxiety. Regarding self-concept, it was believed that whether the test can produce a positive or

negative influence on students' confidence depended on their own opinion about the accuracy of the test results, his/her performance on the test and other individual characteristics. Additionally, the stakes of a test, the frequency with test delivery, and expectations of success or failure on the test can influence a student's learning motivation. It was also found that different types of tests, such as open-book versus closed-book, multiple-choice versus essay questions, influence a student's study practices differently.

Amrein and Berliner (2003) conducted a study on "The effects of High-stakes Testing on Student Motivation and Learning" in which the washback effects of high-stakes testing on students in grades 3-8 of the No Child Left Behind Act were investigated. The research was carried out over eighteen high-stakes testing states in the United States. Through calculating the statistics collected, they explored that the states conducting high school graduation test had higher drop-out rates than those without this test. It means that this kind of tests leads to decrease in students' learning motivation and even increase in dropout rates. To measure effects of high-stakes tests on student learning, archival time-series analysis was applied. Students in these eighteen states took four highly respected measures: the Scholastic Aptitude Test (SAT), American College Test (ACT), Advanced Placement (AP) tests, and the National Assessment of Educational Progress (NAEP) independently. Then the results in different years were compared with national data for each measure. The researchers draw a conclusion that "high-stakes testing policies have resulted in no measurable improvement in student learning" (p. 36).

In their research into the effects of the College English Test (CET) on college

students' English learning in China, Li, Qi & Hoi (2012) investigated students' perceptions of the impact of the CET on their English-learning practices and their affective conditions. A survey was administered to 150 undergraduate students at a university in Beijing. It was found that students perceived the impact of the CET to be pervasive. In particular, most of the respondents indicated that the CET had a greater impact on what they studied than on how they studied. Most of the students surveyed felt the CET had motivated them to make a greater effort to learn English. Many students seemed to be willing to put more effort on the language skills most heavily weighted in the CET. About half of the students reported a higher level of self-efficacy regarding their overall English ability and some specific English skills as a result of taking or preparing for the CET. However, many students also reported experiencing increased pressure and anxiety in relation to learning English.

3. Methodology

3.1. Context and Participants

This study took place at Vietnam National University, Hanoi (VNU), one of the highest-rank universities in Vietnam. As this university requires its students to achieve English proficiency level B1 (Common European Framework of Reference – CEFR), all the students are required to take three English courses consecutively for their first two years. At the end of the last English course (GE3), students take the VSTEP. Students are expected to achieve a certain score on VSTEP in order to receive a bachelor's degree.

In May 2019, 751 VNU students who did not major in English completed a questionnaire that asked them how they felt

about the impact of VSTEP. Of the students who provided demographic data, 149 students were learning GE1, which is the first module in the English program, accounting for 19.84%; 360 students were studying GE2 (the second module) which made up the majority of participants of the study (47.94%); and 242 respondents were taking GE3 as the final module before taking VSTEP (32.22%). The proportion of respondents in the three English modules, though not completely balanced, is also quite diverse, ensuring the representation of all learners in the English program at VNU.

3.2. Questionnaire

A questionnaire was constructed to solicit students' perceptions of the effect of the VSTEP on their English learning. All measurements are made on the Likert-type scale (6 points) with 1 – Strongly disagree, 2 – Disagree, 3 – Slightly disagree, 4 – Slightly Agree, 5 – Agree, 6 – Strongly agree. To ensure validity of the measurement, all items were obtained from previous studies of Putwain & Best (2012) and Mahmoudi (2014) with adjustments to fit the setting of the current study.

There are two main parts in the questionnaire. The first section includes items related to students' perception of the test, namely test difficulty, test familiarity, test importance. The second section elicits information about students' English learning practices in terms of goal setting and study planning, study content and material, study methods and test preparing strategies.

3.3. Data collection and data analysis

Copies of the questionnaire, now rendered in Vietnamese, were distributed to 900 undergraduate students by the researcher of the current study. The purpose and significance of the study were explained to the students, and terminologies were clarified before the students completed the questionnaires. Of 900 copies, 751 were returned to the researcher.

The analytical methods of Cronbach's Alpha, Exploratory Factor Analysis (EFA), Confirmatory Factor Analysis (CFA) and Structural Equation Modeling (SEM) were used for analysis. According to Schumacker & Lomax (1996), structural equation modelling (SEM), which focuses on testing causal processes inherent in theories, represents an important advancement in social work research. Before SEM, measurement error was assessed separately and not explicitly included in tests of theory. With SEM, measurement error is estimated and theoretical parameters are adjusted accordingly.

4. Results

4.1. Descriptive statistics

Test difficulty

The participants of the current study did not attend any official VSTEP at the time of the survey. Their perceptions of the test difficulty were formed through senior students' rumours, teachers' repeated warnings or their experience with mock tests and test-related materials.

Table 1 shows the three items related to students' perceptions of how difficult the VSTEP was, the mean score and standard deviation of each item.

Table 1. Students perception of test difficulty

Item	Mean	Standard deviation
Senior students say that VSTEP is very difficult	4.21	1.259
Teachers say that VSTEP is very difficult	3.66	1.266
After doing mock tests, I feel that VSTEP is very difficult	4.08	1.266

As shown in Table 1, the majority of respondents perceived the difficulty level of the test through senior students' evaluations as this item had the highest mean score of 4.21. Mock tests and test-related materials such as sample tests, past papers of similar tests also played an important role in students' perception of the test difficulty. To the researcher's surprise, teachers seemed not to exert pressure on students by bombarding them with warnings about the difficulty of the test as the third item had the lowest mean score of 3.66.

Test importance

In the questionnaire, there are four

statements that focus on clarifying the importance of the standardized test. These four assessments are divided into two groups: students' judgments about the importance of the test and the importance of the test from teachers' perspective.

Students' judgements about test importance include: (1) If I don't pass the VSTEP, I will be very disappointed; (2) The results of the VSTEP will greatly affect my future work. Teachers' judgements about test importance include: (1) Teachers often remind me of the time to take VSTEP; (2) Teachers often remind me of the consequences of failing VSTEP.

Table 2. Students' perception of test importance

Item	Mean	Standard deviation
Students' judgements about test importance	4.57	1.194
Teachers' judgements about test importance	3.80	1.286

Compared to teachers, the participating students seemingly experienced more anxiety caused by the VSTEP. The item related to students' evaluation of the test significance had a higher mean score than the item linked to teachers' perception with the former receiving 4.57 and the latter 3.80. The students themselves were well aware of the consequential impact that test results might have, but their teachers did not frequently warn them of the detrimental effect that their failure at the test might bring. This finding

corresponds to the previous finding, both of which confirm that teachers acted as an intermediary between the students and the test and they did not stress the difficulty or importance of the test.

Test familiarity

To evaluate students' familiarity with the test, there are three items in the questionnaire, the mean scores of which are shown in the following table.

Table 3. Students' test familiarity

Item	Mean	Standard deviation
I can describe the test format	3.53	1.363
I can name the skills tested in the test	4.16	1.302
I can tell the purpose of implementing VSTEP	3.61	1.266

The results show that students were only confident about the skills tested in the test with an average score of 4.16. Students seemed uncertain about the test format (Mean: 3.53) and the university's purpose of applying the test (Mean: 3.61).

4.2. Inferential statistics

4.2.1. Exploratory Factor Analysis (EFA)

The broad purpose of exploratory factor analysis (EFA) is to summarize data so that

relationships and patterns can be easily interpreted and understood. It is normally used to regroup variables into a limited set of clusters based on shared constructs. After performing EFA, the variables of “Test importance” and “Test difficulty” were merged and renamed as “Pressure from the test”. The factor “Study methods and test preparing strategies” in the suggested model was divided into two new variables, namely “Study methods and test preparing strategies” and “Practice with native speakers”.

4.2.2. Confirmatory Factor Analysis (CFA)

To test the measurement validity, confirmation factor analysis (CFA) was performed. Confirmatory factor analysis (CFA) is a special form of factor

analysis, most commonly used in social research. It is used to test whether measures of a construct (items in the questionnaire) are consistent with a researcher’s understanding of the nature of that construct (or factor).

First, multiple fit indices, including chi-square, degree of freedom (CMIN/DF), goodness of fit (GFI), comparative fit index (CFI) and root mean square error of approximation (RMSEA) were considered. All of our results satisfied the rule of thumb values as illustrated in the following table: Chi-square divided by degree of freedom should be less than 3 ($CMIN/DF \leq 3$) (Carmines & McIver, 1981); GFI and CFI are to be larger than 0.9 (Bentler & Bonett, 1980); RMSEA should be less than 0.08 (Steiger, 1990).

Table 4. The Reliability and Validity of Constructs

Multiple fit indices	Value	“Rule of thumb” values
CMIN/DF	1.851	≤ 3 (Carmines và McIver, 1981)
GFI	0.951	≥ 0.9 (Bentler & Bonett, 1980)
TLI	0.959	≥ 0.9 (Bentler & Bonett, 1980)
CFI	0.970	≥ 0.9 (Bentler & Bonett, 1980)
RMSEA	0.034	≤ 0.08 (Steiger, 1990)

Second, we examined the convergent validity of our measurements through estimation of all items’ construct reliability (CR) and average variance extracted (AVE). As shown in Table 5, all the above indices were satisfied: All CRs (composite reliability)

and AVEs (average variance extracted) are above their cutoff points, that is, 0.8 and 0.5, respectively. Two AVEs were just under 0.5 ($0.494 < 0.50$), but they were still at acceptable level and significant in content value (Nguyễn Đình Thọ & Nguyễn Thị Mai Trang, 2009)

Table 5. Construct validity by Composite reliability and Average variance extracted

Construct	Component	Composite Reliability	Average Variance Extracted
Test factors	Test familiarity	0.743	0.494
	Pressure from the test	0.785	0.513
English learning	Goal setting and planning	0.844	0.581
	Learning content and materials	0.778	0.509
	Learning methods and test preparing strategies	0.585	0.494
	Practice with native speakers	0.760	0.613

Our results indicate that all the constructs in the model have acceptable discriminant validity, and the constructs included in this study are uncorrelated with the others.

4.2.3. Structural equation model (SEM) and hypotheses testing

As all fit indices, including the FI, TLI, CFI and RMSEA satisfied the model fit criteria, they suggest that the whole

structural model proposed in this study is a good fit. The indices include Chi-square=1056.509 ($p < 0.001$), $FI = 0.913 > 0.9$, $TLI = 0.912 > 0.9$, $CFI = 0.924 > 0.9$, and $RMSEA = 0.049 < 0.08$.

These results demonstrate that our proposed model has a significant fit with the obtained data, and all endogenous variables are explainable through exogenous variables included in the framework.

Table 6. The causal relations between constructs in the proposed model

Relation		Estimate	S.E.	C.R.	P-value	
Test pressure	---->	Goal setting and planning	.163	.030	5.400	***
Test pressure	---->	Learning content and materials	.221	.034	6.591	***
Test pressure	---->	Practice with native speakers	.018	.036	.519	.604
Test pressure	---->	Learning methods and test preparing strategies	.261	.050	5.269	***
Test familiarity	---->	Goal setting and planning	.539	.054	9.932	***
Test familiarity	---->	Learning content and materials	.335	.053	6.333	***
Test familiarity	---->	Practice with native speakers	.230	.056	4.086	***
Test familiarity	---->	Study methods and test preparing strategies	.113	.072	1.576	.115

Table 7. Standardized Regression Weights

Relation		Estimate	
Test pressure	---->	Goal setting and planning	0.192
Test pressure	---->	Learning content and materials	0.265
Test pressure	---->	Practice with native speakers	0.022
Test pressure	---->	Learning methods and test preparing strategies	0.286
Test familiarity	---->	Goal setting and planning	0.513
Test familiarity	---->	Learning content and materials	0.324
Test familiarity	---->	Practice with native speakers	0.217
Test familiarity	---->	Learning methods and test preparing strategies	0.100

As can be seen from Table 6, “Pressure from the test” was found to have significant effects on students’ goal setting and planning, students’ selection of learning content and materials, their choice of study methods and exam preparation strategies when P-values are all below 0.05. The weights of these constructs are all positive, respectively

0.163, 0.221 and 0.261, which means that the pressure of the test and these constructs are positively related. The higher the pressure from the test is, the more active the students are in setting goals and planning their study towards VSTEP. Similarly, when the students feel more stressful about succeeding in the test, they tend to choose more test-oriented

materials and are inclined to refuse learning activities that do not directly prepare them for the test. From Table 7, we could find that the test pressure exerted the most influence on “Study methods and exam preparing strategies” (Estimate = 0.286) and the least effect on “Goal setting and study planning” (Estimate = 0.192). However, “Pressure from the test” had no effect on students’ seeking opportunities to practice with foreigners because $p = 0.604 > 0.05$.

In terms of “Test familiarity”, the results also show that the knowledge of the test affected students’ goal setting and study planning, their selection of learning content and materials, their effort to seek opportunities to practice with foreigners (P -values < 0.05). The influence of “Test familiarity” on “Goal setting and study planning” was the largest (Estimate = 0.513) and on Practice with foreigners the smallest (Estimate = 0.217) (see Table 7). It was believed that the more familiar students were with the test, the more specific their study plan was, the content they choose to learn was closer to the test format, and the more active students were in finding opportunities to practice English with foreigners. In contrast, “Test familiarity” did not influence students’ choice of study methods and exam preparing strategies ($p = 0.115 > 0.05$).

5. Discussion

Washback effect of test pressure

From the findings of the current study, it can be seen that students mostly perceived the difficulty and importance of the test through rumors from the senior students and through doing VSTEP coaching materials. In particular, senior students’ evaluations of VSTEP served as the largest factor in constituting the participants’ perception of

VSTEP. The students under study were not subject to the pressure from teachers. This was consistent with Li, Qi & Hoi (2012)’s result as these authors reported a similar trend in China: many students experienced higher pressure and anxiety in relation to learning English when preparing for Chinese English test. This phenomenon derived from the fact that in both China and Vietnam, English-language tests are used as gate-keeping devices for access to general employment and higher education opportunities.

We also found that the pressure of the test affects students’ goal setting and study planning, selection of learning content and materials, choice of study methods and exam preparing strategies. In particular, the pressure from the test had the most influence on study method and exam preparing strategies and the least effect on goal setting and study planning.

The greater the pressure of the test (the more difficult and important it is to students), the more students would proactively set specific goals and plan for their study and choose test-focused materials. The pressure from the test also made students prefer to choose to study at home rather than go to class. When going to class, students did not like to participate in activities that did not help prepare for the test. They also preferred to study alone rather than interact with friends.

Thus, the test clearly makes students more inclined to “study for exams”. A similar trend was also found in a number of previous studies (Karabulut, 2007; Pan, 2009; Tsagari, 2009). The positive effect is that a high-stakes test helps students become more proactive in setting goals and setting a learning path for themselves, which was reported in Huang (2004)’s research in Taiwanese context. In another study, Pan (2009) claimed that Taiwanese students were very supportive of

the English test as a university graduation exam (GEPT) because they thought that the test motivated them to learn English and GEPT certificate helped them to find a job more easily.

However, the results related to the effects of VSTEP on students' choice of learning content and learning methods are quite worrying as students seemed to focus extensively on test coaching. There is no clash between the findings of the current and those of Pan (2009) and Karabulut (2007) when the authors reported that students concentrated on the knowledge and skills that were tested and ignored those which were absent. The high-stakes test in Karabulut (2007)'s research, however, is a university entrance exam in Turkey focusing only on grammar, vocabulary, and reading comprehension, so this effect seemed negative. The test could not improve students' ability to use English in practice. In contrast, VSTEP is a standardized test that fully tests four communicative skills of Listening, Speaking, Reading and Writing. In this case, if students focus only on VSTEP's content and skills, their English communication skills can still be improved. Nevertheless, it is still very important to bear in mind the fact that a test still cannot cover all knowledge and skills that are necessary in life. For example, VSTEP writing section only tests email writing and essay writing skills, it is impossible to test all writing skills for students' future life and work, such as writing reports, making notes, etc. Overreliance on the test could result in students' limited learning experience and inadequate English proficiency.

Regarding learning methods, students tend to prefer self-study at home. This trend reflects a students' lack of confidence in the effectiveness of the English program in

assisting them to pass VSTEP. This case was also mentioned by Pan (2009) in the study of English exam in Taiwanese universities. In his study, Pan (2009) stated that 53% of the surveyed students said they were dissatisfied with English courses offered at the university and they wanted to study at home or at language centers. Students expressed their annoyance and concern as the cost spent on test coaching centers and retaking the university English test was considerably large. However, they still did not choose to study at university because the curriculum was believed not to be effective.

According to the results of this study, the participating students tended to choose activities that were directly related to the test and preferred to study alone instead of participating in interactive activities. This agrees with Li, Qi & Hoi (2012)'s observation that many students learn English for the sake of taking the tests rather than for using the language for real purposes. This is a worrying phenomenon because the nature of learning foreign languages is learning in interaction and using the target language in real-life situations. Teachers need to recognize this trend to design learning activities that both ensure interaction and help students prepare for the test. More importantly, the purpose of communicative activities must be made clear to students so that they know those activities both help them with the test and with their English proficiency.

Washback effect of test familiarity

Regarding test familiarity, students could only name some of the skills that would be tested at the absence of the knowledge about the format and purpose of the test. This finding is fairly surprising since most participants were studying GE2 module

(47.94% of the study subjects) and GE3 (32.33% of the study subjects). These two modules are the final ones before students must take VSTEP; however, they still did not know the format and the purpose of the test. The only thing students were sure of was the skills to be tested in VSTEP, but this was possibly their guess based on the experience with other English standardized tests. Their knowledge of tested skills might be confined to four general language skills of Listening, Speaking, Reading and Writing.

This finding also raises questions about the current English program in general and GE3 in particular. The program and the instructors need to familiarize students with the exam format so that students will be well prepared when they enter the exam room. This is recommended by Hughes (1989) that students need to be well informed of the test format and be well familiarized with tested skills to perform successfully at a high-stakes test. In addition, the purpose of applying a standardized test with all four skills should also be widely disseminated to students so that they understand the direct relationship between the test and their ability to use English in reality. Once the connection between the English program, VSTEP and students' real-life English proficiency are clarified, students' motivation to learn English will be enhanced.

"Test familiarity" was also found to affect students' goal setting and study planning, their selection of learning content and materials, their effort to seek opportunities to practice with foreigners. The influence of "Test familiarity" on "Goal setting and study planning" was the largest and on "Practice with foreigners" the smallest. It was believed that the more familiar students were with the test, the more specific their study plan was, the content they choose to learn was closer to the test format, and the more

active students were in finding opportunities to practice English with foreigners.

The findings of this study suggested that VNU's current English program needs reconsidering and improving. First, to help students pass the standardized test VSTEP, there should be more guide about VSTEP with a specific set of materials such as VSTEP specification, test samples, grammar and vocabulary booklets, especially in the last module GE3. Students consequently will be more familiar with the test and grasp a clearer orientation on how to prepare for it. In that way, the financial burden spent on VSTEP coaching classes and coaching materials can also be reduced. Second, although English competency expected of students is outlined as learning outcomes in the course guide of each English module, this description still needs to be specified in the form of CEFR (Common European Framework for Reference) learning outcomes in terms of grammar, vocabulary, communicative functions to be achieved at each level. Such documents should also be the guide for all learning activities and learning materials. This will help increase the chances of students passing VSTEP and improving their English proficiency up to the expected level. Finally, the teachers should be trained on how to balance between preparing the students for their highly important test and creating communicative activities that encourage and promote English use for real purposes.

6. Conclusion

Based on evidence from the questionnaire survey, VSTEP seems to have had a pervasive impact on the participating students in this study. First, students mostly perceived the difficulty and importance of the test through rumors from the senior students and through doing test-oriented exercises. In particular, senior students' evaluations of VSTEP

served as the largest factor in constituting the participants' perception of VSTEP. The students under study were not subject to the pressure from teachers because the teachers did not often accentuate the difficulty or importance of the test. We also found that the pressure of the test affects students' goal setting and study planning, selection of learning content and materials, choice of study methods and exam preparing strategies. In particular, the pressure from the test has the most influence on study method and exam preparing strategies and the least effect on goal setting and study planning. Second, regarding test familiarity, students could only name some of the skills that would be tested at the absence of the knowledge about the format and purpose of the test. A pattern seemed to emerge whereby test familiarity was also found to affect students' goal setting and study planning, their selection of learning content and materials, their effort to seek opportunities to practice with foreigners. The influence of test familiarity on goal setting and study planning was the largest and on students' efforts to practice with foreigners the smallest.

As VSTEP serves as a gatekeeper for university graduation, students seem to be sensitive to what is assessed. They plan their learning activities and learning content around this high-stakes test and seem to ignore the content and activities which they think are irrelevant. Students also display lack of confidence in the university's English programs. Given the VSTEP's powerful impact on college English education, it is important that the designers and implementers of English program keep reforming the syllabus, teaching materials, teaching methodology and adopt more effective measures in order to encourage students to take more interest in English-language use in real-world contexts in parallel with prepare for successful performance at VSTEP.

References

Vietnamese

Nguyễn Đình Thọ & Nguyễn Thị Mai Trang (2009), *Nghiên cứu khoa học trong quản trị kinh doanh*. Hà Nội: Nhà xuất bản Thống kê.

English

- Amrein, A. L., & Berliner, D. C. (2003). The Effects of High-stakes Testing on Student Motivation and Learning. *Educational Leadership*, 32-38.
- Bentler, P.M. & Bonett, D.G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, 88, 588-606.
- Carmines, E. G., & McIver, J. P. (1981). Analyzing Models with Unobserved Variables: Analysis of Covariance Structures. In G. W. Bohrnstedt, & E. F. Borgatta (Eds.), *Social Measurement: Current Issues* (pp. 65-115). Beverly Hills: Sage Publications, Inc.
- Etten, S. V., Freebern, G. & Pressley, M. (1997). College students' beliefs about exam preparation. *Contemporary Educational Psychology*, 22, 192 – 212.
- Huang, Y. J. (2004). *Exit Requirements worry students?* Liberty Times. Retrieved December, 25, 2007 from <http://www.libertytimes.com.tw/2004/new/oct/17/today-life9.htm>
- Hughes, A. (1993). *Backwash and TOEFL 2000*. Unpublished manuscript. UK: University of Reading.
- Karabulut, A. (2007): *Micro level impacts of foreign language test (university entrance examination) in Turkey: a washback study* (Master's thesis, Iowa State University, Ames, Iowa). Retrieved from <http://lib.dr.iastate.edu/rtd/14884/Date??>
- Kirkland, M. C. (1971). The effect of tests on students and schools. *Review of Educational Research*, 41(4), 303-350.
- Li, H., Qi, Z. & Hoi, K. S. (2012). Students' perceptions of the impact of the college English test. *Language Testing in Asia*, 2(3), 77 – 94.
- Madaus, G. F. (1988). The Distortion of Teaching and Testing High-stakes Testing and Instruction. *Journal of Education*, 65, 29-46.
- Mahmoudi, L. (2014). The washback effects of Iranian national university entrance exam on pre-university English teaching and learning. PhD thesis, Malaya University, Kuala Lumpur, Malaysia.
- Messick, S. (1996). Validity and washback in language testing. *Language Testing* 13(4), 241-56.
- Minarechova, M. (2012). Negative impacts of high-stake tests. *Journal of Pedagogy*, 3(1), 82-100.
- Orfield, G. & Wald, J. (2000). Testing, testing: The high-stakes testing mania hurts poor and minority students the most. *The Nation*, 270(22), 38 – 40.

- Pan, Y. (2009). Test impact: English certification exit requirements in Taiwan. *TEFLIN Journal*, 20 (2), 119 – 134.
- Putwain, D. & Best, N. (2012). Fear appeals in the primary classroom: Effects on test anxiety and test grade. *Learning and Individual Differences*, 21, 580 -584.
- Rea-Dickins, P. (1997). So, why do we need relationships with stakeholders in language testing? A view from the UK. *Language Testing*, 14(3), 304–314.
- Schumacker, E. R., & Lomax, G. R. (1996). A beginner's guide to structural equation modeling. Mahwah, NJ: Erlbaum.
- Shohamy, E. (2001). *The power of tests*. London: Longman/ Pearson.
- Steiger, J.H. (1990). Structural Model Evaluation and Modification: An Interval Estimation Approach. *Multivariate Behavioral Research*, 25(2), 173–180.
- Tsagari, D. (2009). *The Complexity of Test Washback: An Empirical Study*. Language Testing and Evaluation Series, Grotjahn, R. & G. Sigott (eds). Frankfurt am Main: Peter Lang GmbH.
- Wall, D. & Alderson, J.C. (1993). Examining washback: the Sri Lanka impact study. *Language Testing*, 10(1), 41-69.
- Wall, D. (1997). Impact and washback in language testing, In C. C. & D. Corson (Eds.), *Encyclopedia of language and education. Volume 7: Language Testing and Assessment* (pp 291 – 302). Dordrecht, the Netherlands: Kluwer.

CẢM NHẬN CỦA NGƯỜI HỌC VỀ BÀI THI ẢNH HƯỞNG TỚI HOẠT ĐỘNG HỌC TIẾNG ANH NHƯ THẾ NÀO? MỘT NGHIÊN CỨU TẠI ĐẠI HỌC QUỐC GIA HÀ NỘI

Nguyễn Thúy Lan¹, Nguyễn Thúy Nga²

1. Phòng Đào tạo, Đại học Ngoại ngữ, ĐHQG Hà Nội, Phạm Văn Đồng, Cầu Giấy, Hà Nội, Việt Nam
2. Đại học Giáo dục, ĐHQG Hà Nội, 144 Xuân Thủy, Cầu Giấy, Hà Nội, Việt Nam

Tóm tắt: Hiện tượng ảnh hưởng của bài thi tới hoạt động học tập từ lâu đã nhận được nhiều sự quan tâm của các nhà nghiên cứu. Trong các yếu tố ảnh hưởng đến hoạt động học ngoại ngữ, cảm nhận của người học đóng vai trò vô cùng quan trọng, đặc biệt khi người học phải vượt qua một bài thi quan trọng như bài thi xét tốt nghiệp đại học. Nghiên cứu này cố gắng tìm ra câu trả lời cho câu hỏi: Liệu cảm nhận của người học về độ khó, tầm quan trọng và độ quen thuộc với bài thi ảnh hưởng thế nào tới hoạt động học tiếng Anh của họ? Nghiên cứu tiến hành khảo sát 751 sinh viên không thuộc chuyên ngành Ngôn ngữ Anh tại Đại học Quốc gia Hà Nội. Các phương pháp phân tích như Phân tích nhân tố khám phá (EFA), Phân tích nhân tố khẳng định (CFA) và Mô hình phương trình cấu trúc (SEM) được sử dụng để phân tích số liệu thu được từ bảng hỏi. Kết quả thực nghiệm cho thấy VSTEP có tác động tương đối lớn tới sinh viên tham gia nghiên cứu. Cảm nhận của sinh viên về độ khó và tầm quan trọng của bài thi VSTEP chủ yếu được hình thành qua sự chia sẻ của các sinh viên khóa trên mà không bắt nguồn từ giảng viên. Áp lực từ bài thi và mức độ quen thuộc với bài thi tỉ lệ thuận với sự chủ động của sinh viên trong việc xác định mục tiêu và lập kế hoạch học tập; hai yếu tố này cũng ảnh hưởng đến việc sinh viên lựa chọn nội dung và tài liệu học tập. Trong khi đó, áp lực từ bài thi không ảnh hưởng đến nỗ lực của sinh viên trong việc tìm kiếm cơ hội thực hành với người nước ngoài, và mức độ quen thuộc của bài thi không ảnh hưởng đến phương pháp học tập và chiến lược luyện thi. Kết quả cũng cho thấy sinh viên có xu hướng lựa chọn các hoạt động và nội dung học tập theo định hướng của bài thi VSTEP mà không quan tâm đến các nội dung không có trong bài thi.

Từ khóa: cảm nhận của người học, bài thi quan trọng, ảnh hưởng đối ngược, định hướng từ bài thi, SEM.