

## Washback of English national exams at ninth-grade level in Thailand and Indonesia

Wachyu Sundayana<sup>1\*</sup>, Panusak Meekaeo<sup>2</sup>, Pupung Purnawarman<sup>3</sup>, and Didi Sukyadi<sup>4</sup>

English Education Department, Faculty of Language and Literature Education, Universitas Pendidikan Indonesia, Bandung, West Java, Indonesia<sup>1,3,4</sup>

English Education Department, School of Postgraduate, Universitas Pendidikan Indonesia, Bandung, West Java, Indonesia<sup>2</sup>

### ABSTRACT

Washback refers to influences of testing on teaching and learning. In Thailand and Indonesia, washback of the Ordinary National Educational Test (O-NET) and *Ujian Nasional* (UN) inevitably occurs on teaching and learning in classrooms at every level. This present study aims to explore and compare the washback effects of the O-NET and UN on English language learning as perceived by Thai and Indonesian ninth-grade students. It is a multi-case study (Thailand case and Indonesia case) by using triangulation design as the research design. The questionnaires concerning washback effects of national exams on English language learning were distributed to 200 ninth-graders in the two cases. In addition, six students from each case were interviewed. The results reveal that in both cases, the participants focus to learn on contents and skills that were likely to appear in the national exams. The participants learned English harder to perform well in the tests rather than to improve their English ability. Moreover, the students had high anxiety during the test preparations and feared for low O-NET and UN scores. The results of the study contribute to future washback study and improvement of English language teaching and learning at ninth-grade in Thailand and Indonesia.

**Keywords:** washback effects; Ordinary National Educational Test (O-Net); *Ujian Nasional* (UN); English language teaching and learning

**First Received:**

1 April 2017

**Revised:**

9 January 2018

**Accepted:**

9 March 2018

**Final Proof Received:**

28 May 2018

**Published:**

31 May 2018

### How to cite (in APA style):

Sundayana, W., Meekaeo, P., Purnawarman, P., & Sukyadi, D. (2018). Washback of English national exams at ninth-grade level in Thailand and Indonesia. *Indonesian Journal of Applied Linguistics*, 8(1), pp. 167-176. doi: 10.17509/ijal.v8i1.11478

### INTRODUCTION

Washback (Alderson & Wall, 1993) or backwash (Biggs, 1995 in Cheng & Curtis, 2004) (henceforth washback) refers to influences of testing on teaching and learning processes (Bailey, 1996, 1999; Cheng, 2000; Cheng & Curtis, 2004). In line with this, Wall and Alderson (1993: p. 41) and Bailey (1999) claim that washback of tests can have powerful influence, both positive and negative, on classroom settings. In high-stakes standardized testing, washback refers to an approach to gather information from students or schools at a certain level (The Greenwood Dictionary of Education, 2012). Sullivan (2009) adds that failing to pass a test, or obtain a low score, could have detrimental consequences for students.

High-stakes standardized testing in Thailand is associated with the Ordinary National Educational Test (henceforth O-NET). The O-NET is the national examination that has been used to measure Grade 6, 9, and 12 students' attainment of graduate learning outcomes at the end of their final semester nationwide (National Institute of Educational Testing Service [NIETS], 2012). The purposes of the test are to test students' academic knowledge and thinking ability, to provide information to schools about improvement of teaching and learning activities, and to evaluate the quality of education at the national level (NIETS, 2012).

Similarly, in Indonesia, *Ujian Nasional* (literally means National Examination) has been used for similar reasons as the O-NET in Thailand. The *Ujian Nasional*

\* Corresponding author:  
Email: swachyu@upi.edu

(henceforth UN) aims to evaluate Indonesian students' attainment of *Standar Kompetensi Lulusan* (SKL – Graduate Learning Outcomes) at the end of each educational level (Sukyadi & Mardiani, 2011). The SKL covers knowledge competence, skill competence (including thinking skills), and attitude competence. The UN is emphasized on knowledge and skill competencies. Since 2015, the UN result is not used as a single determiner of a student graduation decision. However, the UN is still considered as a high-stakes testing because the result still will be used to map out the quality of Indonesian education, to select students at higher levels of education, and to plan appropriate actions and funding schemes to support schools in order to improve the quality of education at schools and district levels (Sulistyo, 2009; Saukah, 2015).

Washback effects of the O-NET and UN, therefore, inevitably occur in the teaching and learning processes at every tested level. Bailey (1999) says that tests have a powerful influence on learners who are preparing for such test, and on teachers who try to help students to prepare. Students of the ninth-grade level have to take English as one of the tested subjects. The washback effects of the O-NET and UN, consequently, play a vital role in classroom settings that affects teaching and learning activities. Nevertheless, there is limited number of researches about washback effects of the O-NET and UN on English language learning at ninth-grade level. Currently, the washback studies generally focus on the effects of tests on teaching (Pan & Newfields, 2012, in Akpınar & Cakildere, 2013) rather than on learning. The washback of tests on learning should be explored (Wall, 2000; Watanabe, 2004) also because students are affected by those tests directly (Pan & Newfields, 2012, in Akpınar & Cakildere, 2013).

Studies on test anxiety, for example, indicate that test has significant effect on learning process (Aydin, 2009). Students with high anxiety do not perform better in foreign language learning. In the context of national examination, there is only a limited number of researches concerning washback effects of the high-stake tests, like O-NET and UN, on English language teaching and learning at ninth-grade level. Therefore, the followings studies are just relevant to the focus of the present study. To fill the gap, this study aims to explore and to compare the washback effects of the O-NET and UN on English language learning as perceived by Thai and Indonesia ninth-graders. The results of the study is expected to contribute to the understanding and the improvement of English language teaching and learning at ninth-grade level in Thailand and Indonesia.

### **Definitions of Washback**

Washback refers to the impact of a test on teaching and learning (Bailey, 1996, 1999; Cheng, 2000; Cheng & Curtis, 2004). Additionally, it can be the extent to which tests influence language teachers and learners to do what 'they would not otherwise necessarily do' (Alderson & Wall, 1993). In addition, Cheng and Curtis

(2004) list some scholars' definitions as follows. Washback (Alderson & Wall, 1993) or *backwash* (Biggs, 1995 in Cheng & Curtis, 2004) refers to the impact of testing on teaching and learning. Popham (1993) associates washback as *measurement-driven instruction*. It refers to the notion that tests should drive teaching and then bolster learning. Furthermore, Shepard (1990, 1991, 1992, 1993) links washback to *curriculum alignment* (Shepard, 1990, 1991, 1992, 1993). It is focused on the connection between curriculum-based teaching practice and the testing.

In addition, other terms that are connected to washback are *systemic validity* (Frederiksen & Collins, 1989), *consequential validity* (Messick, 1989, 1992, 1994, 1996), and *test impact* (Bachman & Palmer, 1996). They are related to the integration of tests into educational system. Nevertheless, Bailey (1996), different from others, defines washback as the influence of testing on teaching and learning processes. This is widely held to exist and to be important. In fact, it is relatively little numbers of empirical researches that document its exact nature or mechanisms by which it works.

### **Negative and Positive Washback**

Tests can provide either negative or positive effect to stakeholders. Negative washback refers to unexpected and detrimental effects of testing. For example, teachers may focus too heavily on test preparation rather than other learning activities (Center for Applied Linguistics[CAL], 2016). Similarly, Taylor (2005) explains that negative washback arise when format or content of a test focuses on specific language ability; and so constrains the teaching or learning context.

Furthermore, Akpınar and Cakildere (2013) say that teachers and students would not improve other skills but the tested skills in the exams. Some skills would be ignored while other skills would be emphasized. Besides, Cheng (2003) notes that a high-stakes test can drive superficial changes in teaching practice. Even though, test encourages teachers to change textbooks and to emphasize new skills, the teachers intend to maintain their teaching methodology. On the other hand, positive washback refers to expected test impacts, e.g. a test may encourage students to study more (CAL, 2016). Cheng (2003) asserts that positive washback scenario is related closely to 'measurement-driven instruction' in general education. It means that teachers and learners have positive attitude towards examinations and work to achieve their objectives willingly and collaboratively. In other words, it is a result when a testing encourages good teaching and learning practice. For instance, an oral proficiency test is introduced in the expectation that it will promote the teaching of speaking skills (Taylor, 2005). Akpınar and Cakildere (2013) also support this. They insist that students deeply focus on what will be tested in examinations. Benjamin and Pashler (2015) guarantee that a well-designed test can promote good learning in a number of ways and standardized tests are critical to

assess students' learning and a nation's progress and to ensure international competitiveness of the graduates.

#### **Ordinary National Educational Test (O-NET)**

O-NET is a standard-based achievement test (NIETS, 2012). The test is intended to be administered to assess Thai students' academic ability in their final semester of three grades: Year 6, Year 9, and Year 12 (NIETS, 2015). The main purposes aim to test the knowledge and thinking ability of Grade 6, 9 and 12 students according to the Basic Education Core Curriculum B.E 2551 (A.D. 2008). In addition, it aims to provide information to schools to improve their teaching and learning activities, and to evaluate the quality of education at national level (*ibid.*). The O-NET comprises of five major subjects of the National Education Curriculum (NIETS, 2015), including (1) Thai Language; (2) Mathematics; (3) Science; (4) Social Studies, Religions, and Cultures; and (5) Foreign Languages

The test contents are created based on the Basic Educational Curriculum B.E. 2551 (A.D. 2008) (NEITS, 2012). It is administered by The National Institute of Educational Testing Service (NIETS – a public organization).

#### **Ujian Nasional (UN)**

UN is a standard educational assessment of primary and secondary education in Indonesia. The UN is administered by the National Standard of Education (BSNP – *Badan Standar Nasional Pendidikan*). Sukyadi and Mardiani (2011) explains that Indonesia has administered the UN for its students since 2005 (Government Regulation number 19 of 2005). The test requires sixth, ninth, and twelfth graders to take the national exam at the end of the junior and senior high school (Grade 9 and 12) as a minimum graduation requirement (UNESCO, 2010). The purposes of the UN are to measure and to evaluate the Indonesian students' competences after the process of teaching and learning at the end of each educational level (i.e. grade six, nine, and twelve) (Sukyadi & Mardiani, 2011). The test result of the UN is a part of a system implemented to improve the quality of national education (Sulistyo, 2009). It is expected to be utilized by school principals, teachers, and stakeholders, as well as policy and decision makers at the national level as an instrument in enhancing the Indonesian education (Hamied, 2010).

#### **Previous Studies**

As previously mentioned, there is only a limited number of researches concerning washback effects of the O-NET and UN on English language teaching and learning at ninth-grade level. Therefore, the followings studies are just relevant to the focus of the present study.

#### **In Thailand**

Pukmai (2009) studied the washback effects of the O-NET on English language testing in Srisaket Province, Thailand. He intends to investigate the relationship between teacher-made test scores and the O-NET scores in the academic year 2006. The results reveal that the

correlation score between those tests is low. The scores from teacher-based assessment are twice higher than the O-NET scores. This can be caused by the fact that some teachers might have upgraded or inflated their students' scores for some reason. For instance, it can be intended to increase students' possibilities to pass university admission, while the O-NET scores reveal their real ability of English based on the national curriculum.

Lunrasri and Gajaseni (2014) studied grade nine students' perceptions and opinions towards the washback effects of the O-NET on English language learning in the academic year 2013 in Chachoengsao Province, Thailand. It found that the test influenced the students to focus their learning heavily on the test contents. They were pushed to study harder to develop their language ability and test performance. Most of them memorized the learning content and feared the low O-NET scores.

#### **In Indonesia**

Husnawati (2004, cited in Sulistyo, 2009) criticizes the UN as a serious flaw. In other words, UN might lead students to mechanistic rote learning, but with the sole purpose in response to the test while other potentials are ignored. Furthermore, national exam do not stimulate students' logical thinking, innovative thoughts, and emotional survival (Santoso, 2004, cited in Sulistyo, 2009). Besides, not only viewed from students' factor, Santoso (*ibid.*) asserts that the test items in the test are oriented to academic pedagogical perspectives, building a possible character for the students to be thoughtless to their environment. Furthermore, Zubaidi (2014) and Virgo (2016) state that teachers sometimes coach or teach their students to face tests. As a result, students tend to answer questions directly instead of elaborating their answer by exposing English as a language. Moreover, the UN only focuses on the passive aspects of English and it forces the students to memorize answers (Zubaidi, 2014).

Musthofa (2004, cited in Sulistyo, 2009) comments that everyone needs to support the implementation of the UN for some reasons. First, the UN constitutes a quality of education and influences the country to get rid of the image of the low quality of national education. Second, the UN is a motivating force for students to learn. Third, teachers will be encouraged to improve their teaching performance. Fourth, the last, the UN can bolster the competitiveness of Indonesians in the challenges in the global era through emphasizing Indonesian and English languages and Math testing.

#### **In other countries**

Akpinar and Cakildere (2013) studied the washback effects of high-stakes language tests (KPDS and ÜDS) on language development of the learners in Turkey. The study focused on receptive and productive language skills of academic personnel studying at Nevúehir University, Turkey. It found that the tests had positive washback effects only on tested skills (reading skills)

while writing, listening and speaking skills were totally neglected by students because these skills were not tested.

Kenedy and Lui (2013) studied washback effects from a high-stakes test for university applicants in Beijing, China (Beijing Matriculation English Test: BMET). The aspect under investigation was students' and teachers' perceptions of the role of final-year secondary school English classes in preparing for the test and improving English skills. Results showed that both the students and teacher perceive that a major role of the final year English class is to help students to prepare for the test. Furthermore, the improvement of English skills is of lesser importance.

To conclude, both negative and positive impacts occur among high-stakes tests. Based on some related studies above, washback inevitably happened. This study will focus on washback effect of the O-NET and UN ninth-grade on English learning in Thailand and Indonesia.

## **METHOD**

This study was a multi-case study, i.e. Thailand case and Indonesia case. It employed triangulation or convergent parallel design (Creswell, 2009). It aims to explore and compare the washback effects of the O-NET and UN on English language learning as perceived by ninth-grade Thai and Indonesian students. The quantitative data were collected by using questionnaires. Besides, semi-structured interviews were conducted to gather qualitative data. This study integrated both data into overall interpretation. Then cross-case analysis was employed to integrate the results from each triangulation or convergent study across the two cases (Stake, 2006).

## **Subjects**

The subjects of the present study are 200 Thai and Indonesian ninth-graders. They are students at two inner-city secondary schools in Nangrong City, Buriram, Thailand and two inner-city junior high schools in Bandung City, West Java, Indonesia in the academic year 2016. There were 38 males and 62 females in Thailand case, and 40 males and 60 females in Indonesia case. All of them were asked to complete questionnaires. In addition, six students from each case were randomly selected to be involved in in-depth interviews.

## **Instruments**

There were two kinds of instruments, namely questionnaire and semi-structured interviews. The instruments were adapted from Lunrasri and Gajasenri (2014).

The questionnaire was constructed based on the framework of Shih's washback model of learning (Shih, 2007). There are five learning areas in the washback model of learning including content of learning, total time of learning, learning strategies, learning

motivation, and test anxiety. For the interview, the questions are related to washback effects of the O-NET on English learning. It was also based on areas of Shih's washback model of learning (Shih, 2007).

## **Data analysis**

The data obtained from the questionnaires were analyzed by using SPSS program version 16 to calculate descriptive statistics including arithmetic means, standard deviation, and independent sample t-test.

The interpretation of mean scores to analyze the washback effects on learning are as follows.

4.21 – 5.00 = always did it.

3.41 – 4.20 = often did it.

2.61 – 3.40 = sometimes did it.

1.81 – 2.60 = seldom did it.

1.00 – 1.80 = never did it

Moreover, data from the interviews were analyzed. The two kinds of data were compared, related, and interpreted. Then cross-case analysis was employed to integrate the results from two cases (Stake, 2006).

## **FINDING AND DISCUSSION**

The results shown in Table 1 and Table 2 were categorized into five areas based on Shih's washback model of learning (Shih, 2007). It comprises of content of learning, total time of learning, learning strategies, learning motivation, and test anxiety.

### **Content of learning**

Based on the data in Table 1, it can be inferred that there are no significant differences at .05 level between ninth-grade Thai and Indonesian students' perceptions on learning contents and skills of English ( $t = .61$ ,  $p = .54$ ). However, there are significant differences at .05 level regarding the focus of learning communicative English language skills ( $t = 5.76$ ,  $p = .00$ ) and learning some parts in the English textbooks even though they are not likely to appear in the O-NET and UN ( $t = 3.10$ ,  $p = .00$ ).

In both cases, either Thai or Indonesian students, they often focus learning on contents and skills that were likely to appear in the O-NET and UN ( $M_3 = 3.68$ ,  $SD_3 = .80$ ). In every school, O-NET and UN preparation programs, which all the students have to participate, were held before the tests. The focuses of the preparations are the tested contents and skills that were likely to appear in the O-NET or UN. The programs are conducted in the 2<sup>nd</sup> semester around three or four months before the test date.

Considering students' perceptions about learning communicative skills and using English textbooks, Thai students often relied on English textbooks. In addition, they learn to communicate based on the textbooks. However, their teachers always use L1 in the classroom so that they require more English exposure from the teachers. In contrast, Indonesian

students learn from various learning materials rather than the textbooks. Interestingly, English was usually

spoken in the classrooms while L1 was used when complex explanations were needed.

Table 1. Means, standard deviation, and independent sample t-test of content of learning and total time of learning

	Perceptions of Students				t-test		
	Thai Students (n=100)		Indonesian Students (n=100)		t	df	p
	M	SD	M	SD			
<b>Content of Learning</b>							
1. Focusing learning on the contents and skills of English that are likely to appear in the National Exam.	3.72	.79	3.65	.82	.61	198	.54
2. Focusing learning on communicative English language skills.	3.98	.81	3.33	.77	5.76	198	.00
3. Focusing learning on some parts in the English textbook even though they are not likely to appear in the National Exam.	3.53	.96	3.11	.95	3.10	198	.00
<b>Total Time of Learning</b>							
1. Spending time in the evenings or weekends for the National Exam preparation provided by schools or in tutorial schools.	2.78	1.19	3.30	.94	-3.41	198	.00
2. Spending time in the evenings or weekends to improve their English proficiency e.g., watching English movies, listening to English songs and reading English books.	3.19	1.14	3.65	1.03	-2.97	198	.00
3. Spending time practicing previous National Exam test or reviewing grammar and vocabulary in classrooms.	2.89	1.08	3.10	.90	-1.48	198	.13
4. Spending time practicing communicative English language skills in classrooms.	3.46	1.00	2.82	.65	5.31	198	.00

\*p < .05

Table 2 Mean, standard deviation and independent sample t-test of learning strategies, learning motivation and test anxiety

	Perceptions of Students				t-test		
	Thai Students (n=100)		Indonesian Students (n=100)		t	df	p
	M	SD	M	SD			
<b>Learning Strategies</b>							
1. Learning test-taking strategies for English language tests.	3.38	1.02	2.92	1.03	3.16	198	.00
2. Using rote-memorization to prepare for the National Exam.	3.56	1.03	3.30	.91	1.88	198	.06
<b>Learning Motivation</b>							
1. Studying harder in English in order to develop their ability to use language.	3.28	1.00	3.62	.87	-2.55	198	.01
2. Studying harder in English to get high National Exam scores.	3.38	1.07	3.93	.85	-4.01	198	.00
<b>Test Anxiety</b>							
1. Feeling anxious while preparing for the National Exam.	3.69	1.25	3.37	1.24	1.81	198	.07
2. Fear for the poor National Exam results in English.	4.03	1.16	3.47	1.26	3.25	198	.00

\*p < .05

**Total time of learning**

There are no significant differences at .05 level between ninth-grade Thai and Indonesian students' perceptions in terms of spending time practicing previous National Exam test or reviewing grammar and vocabulary in classrooms (t = -1.48, p = .13). However, there are significant differences at .05 level between ninth-grade Thai and Indonesian students' perceptions concerning spending time for the National Exam preparation provided by schools or in tutorial schools (t = -3.41, p = .00), spending time in the evenings or weekends to improve their English proficiency (t = -2.97, p = .00),

and spending time practicing communicative English language skills in classrooms (t = 5.31, p = .00).

As shown in Table 1, Thai and Indonesian students perceive that sometimes they spent time practicing previous O-NET and UN tests or reviewing grammar and vocabulary in classrooms (M<sub>1</sub> = 2.89, SD<sub>1</sub> = 1.08 and M<sub>2</sub> = 3.10, SD<sub>2</sub> = .90, respectively). However, Indonesian ninth-graders have higher mean scores than Thai students. Moreover, in both cases, the students practiced previous national exam tests and reviewed some language points in the classroom, but it was not all the time because they had O-NET and UN preparations to deal with the tests specifically. At the test preparation

periods, the students in both cases learned hard, and the focuses were on the O-NET and UN tested skills and contents only.

There were significant differences at .05 level between ninth-grade Thai and Indonesian students' perceptions in some items. Firstly, regarding spending time for the O-NET and UN preparations provided by schools or in tutorial schools, both Thai and Indonesian ninth-grade students did it sometimes ( $M_1 = 2.78$ ,  $SD_1 = 1.19$  and  $M_2 = 3.30$ ,  $SD_2 = 1.10$ , respectively). However, ninth-grade Indonesian students had higher mean scores than ninth-grade Thai students. It indicated that Indonesian students attended tutorial schools more than Thai students. Similarly, the purposes of attending the tutorial schools were to improve the school assessment and also the O-NET and UN tested subjects' scores.

In regular classes, ninth-grade Thai students learned based on textbooks, but Indonesian students also learned from other materials besides the textbooks. Previous O-NET and UN tests and language points were not emphasized at this time. In all cases, they had the test preparations to deal with the O-NET and UN specifically. At the test preparations, the students learned hard, and the focus was on the O-NET or UN contents only. The test preparations were conducted in the 2<sup>nd</sup> semester. In Indonesia case, the UN preparations were held about three months before the test (including a month of intensive UN preparations).

During those three months, besides reviewing and practicing previous UN tests, there were six try-out tests that ninth-grade Indonesian students had to take to evaluate their improvement. In contrast, in Thailand case, the O-NET preparations were held only a month, including a week of intensive O-NET preparation before the test. Moreover, there was only one try-out test called Pre-O-NET test.

Secondly, while Thai students spent some time to improve their English at weekends or evenings sometimes ( $M_1 = 3.19$ ,  $SD_1 = 1.14$ ), Indonesian students often did it ( $M_2 = 3.65$ ,  $SD_2 = 1.03$ ). Thai students did not practice English much outside the classroom. They relied on dictionary or the internet to help them understand English. Most Indonesian students often practiced English after class at home. They listened to music, watched movies, and read English books. However, it depended on individuals.

Finally, Thai students often spent time practicing communicative English language skills in classrooms ( $M_1 = 3.46$ ,  $SD_1 = 1.00$ ), but Indonesian students only did it sometimes ( $M_2 = 2.82$ ,  $SD_2 = .65$ ). In Thailand case, interestingly, though ninth-grade Thai students perceived they often practice communicative English in the classroom, L1 was always used in the classroom rather than English. On the other hand, Indonesian students used English as the main language in their classrooms. They sometimes focused learning on communicative English skills. However, they were encouraged to use English in the classroom. L1 was

sometimes used when the students learned some difficult lessons, and they needed deeper explanations.

### **Learning strategies**

There were no significant differences at .05 level between ninth-grade Thai and Indonesian students' perceptions on using rote memorization to prepare the O-NET and UN ( $t = 1.88$ ,  $p = .06$ ). However, there were significant differences at .05 level between ninth-grade Thai and Indonesian students' perceptions in terms of learning test-taking strategies for English language tests ( $t = 3.16$ ,  $p = .00$ ).

As it can be seen from the Table 2, while most Thai students perceived that they often used the rote-memorization technique to prepare the O-NET ( $M_1 = 3.56$ ,  $SD_1 = 1.03$ ), Indonesian students perceived they used that technique to prepare the UN sometimes ( $M_2 = 3.30$ ,  $SD_2 = .91$ ). Ninth-grade Thai students tried to memorize vocabulary because they thought that if they remembered more words, they would get more scores. Likewise, ninth-grade Indonesian students also used this technique and read English books to help them to expand their vocabulary range.

There were significant differences at .05 level between ninth-grade Thai and Indonesian students' perceptions regarding learning test-taking strategies for English language tests. The levels of frequency of Thai and Indonesian students were the same. They perceived that they sometimes learned test-taking strategies for English language tests ( $M_1 = 3.38$ ,  $SD_1 = 1.02$  and  $M_2 = 2.92$ ,  $SD_2 = 1.03$ , respectively). Ninth-grade Thai students learned some strategies to deal with the O-NET, for example, the concordance between questions and choices, word guess and skimming reading. Similarly, ninth-grade Indonesian students learned some strategies to deal with the UN such as checking an answer from a question, guessing the meanings of words from their contexts, and trying to understand each question completely before making any decisions about the answers.

### **Learning motivation**

There were significant differences at .05 level between ninth-grade Thai and Indonesian students' perceptions concerning studying harder in English in order to develop their ability to use the language ( $t = -2.55$ ,  $p = .01$ ) and to get high National Exam scores ( $t = -4.01$ ,  $p = .00$ ).

While, most ninth-grade Thai students sometimes learn English harder to improve their ability to use the language ( $M_1 = 3.28$ ,  $SD_1 = 1.00$ ) and to get high O-NET scores ( $M_1 = 3.38$ ,  $SD_1 = 1.07$ ), most ninth-grade Indonesian students often did it to improve their English ( $M_2 = 3.62$ ,  $SD_2 = .87$ ) and to get high UN scores ( $M_2 = 3.93$ ,  $SD_2 = .85$ ). According to the mean scores, ninth-grade Thai students did not emphasize learning English to improve their language ability more than to get high O-NET scores or vice versa, while ninth-grade Indonesian students gave the precedence of the UN over the improvement of English language ability. Besides,

some ninth-grade Thai students did not stress the importance of English much. Therefore, they did not study beyond English lessons at school. In contrast, some ninth-grade Thai students worried about the school assessment and O-NET scores so that they studied English more after school at tutorial schools. However, the focus was on getting high tests' scores rather than to develop English language ability. In terms of Indonesia case, the UN influenced most ninth-grade Indonesian students to learn English harder. Most of them learned English in order to improve their English ability and to get high UN scores. However, it depended on individuals whether or not they learned English harder.

### **Test anxiety**

There were no significant differences at .05 level between ninth-grade Thai and Indonesian students' perceptions regarding anxiety while preparing for the National Exam ( $t = 1.81, p = .07$ ). However, there were significant differences at .05 level between ninth-grade Thai and Indonesian students' perceptions concerning the fear for the poor national exam results in English ( $t = 1.24, p = .00$ ).

Most ninth-grade Thai students perceived that they often felt anxious while preparing for the O-NET ( $M_1 = 3.69, SD_1 = 1.25$ ). In contrast, ninth-grade Indonesian students perceived that they felt the anxiety from the UN preparation sometimes ( $M_2 = 3.37, SD_2 = 1.24$ ). In Thailand case, the O-NET preparations significantly influenced ninth-grade Thai students' feeling towards the O-NET. The preparations made them felt anxious about reviewed contents and concerned about the difficulty of the test and poor scores. Even though they felt stressful during the preparations, but, they perceived the preparations were useful. They thought the preparations would help them to be confident and ready for the O-NET test. Similarly, in Indonesia case, most ninth-grade Indonesian students were highly stressful because there were enormous contents that need to be reviewed. Furthermore, there were six try-out tests. It was once in two weeks. Therefore, they felt overwhelming and stressful. However, they thought that the UN preparations helped them to be ready for the UN.

There were significant differences at .05 level between ninth-grade Thai and Indonesian students' perceptions on the subject of fear for the poor national exam results in English. Among ninth-grade Thai and Indonesian students, they shared the same level of frequency. They perceived that they often feared for the poor national exam results in English ( $M_1 = 4.03, SD_1 = 1.16$  and  $M_2 = 3.47, SD_2 = 1.26$ , respectively). However, the mean score of ninth-grade Thai students was higher than ninth-grade Indonesian students'. It illustrated that most ninth-grade Thai students had a higher concern about the poor results in English than Indonesian students'. Nevertheless, in both cases, the students feared that they would obtain low scores in English.

### **Discussion**

The results of the present study manifested that the O-NET and UN had both negative and positive impacts to ninth-grade Thai and Indonesian students' English learning.

In both cases, the contents and skills of learning were related to what were likely to appear in the national exams (O-NET and UN). Bailey (1999) said that tests have a powerful influence on learners who are preparing for such test, and on the teachers who try to help them to prepare. Also schools set extra classes to deal with the tests. These results were confirmed by Badan Standar Nasional Pendidikan (henceforth, BSNP) (2009) that some Indonesian schools added more additional hours for the tested subjects in the UN because they thought the time allotment for each subject in the curriculum was not adequate which was similar to Thai schools. The O-NET and UN preparations all started in the 2<sup>nd</sup> semester before the tests in order to help students to prepare the tests. The periods of the preparations varied from school to school. In Thailand case, all the schools spent about a month before the O-NET while in Indonesia case, it took about three months. The preparation periods covered regular and extra classes. Moreover, the fact remains that the students in both cases learned hard in the preparation sessions to the O-NET and UN.

Besides, mostly, the students' learning in both cases relied on what their teachers provided. As a result, assumingly, the teacher emphasized their instruction on what the test tested. In line with this, Vernon (1956: 166, in Alderson & Wall, 1993) negatively claimed that tests 'distort the curriculum' because teachers tended to teach what would appear in the tests and ignored other contents. Also, "what is assessed becomes what is valued, which becomes what is taught" (McEwan, 1995: 42, in Cheng, 2000). However, since the O-NET and UN were the results of Thai and Indonesian curriculum, respectively, therefore, the test drove the instruction and then learning. It is what Popham (1993) called it '*Measurement-driven instruction.*'

What's more, there was only one try-out test, called Pre-O-NET test, for ninth-grade Thai students, but six times for Indonesian students. It indicated that among ninth-grade Indonesian students, they had high anxiety even though the quantitative data results showed that they only felt anxious sometimes ( $M_2 = 3.37, SD_2 = 1.24$ ). It was confirmed by the qualitative data results that ninth-grade Indonesian students had highly stressful because there were enormous contents that need to be reviewed. This finding was identical with a study conducted by Lunrasri and Gajasen (2014). They examined the impacts of the O-NET at grade 9 level in Chacheongsao Province in Thailand. It found that the students had high level of anxiety during the O-NET preparation. Also ninth-grade Thai and Indonesian students highly feared for the poor O-NET and UN scores. The mean scores between them were the same level of frequency ( $M_1 = 4.03, SD_1 = 1.16$  and  $M_2 = 3.47, SD_2 = 1.26$ , respectively). It illustrated that the O-

NET and UN scores were significantly important to them. When considering the usage of the O-NET and UN scores, they support the claims above. In Thailand case, the O-NET scores will be applied to students' GPA and admission into the upper secondary level. Similarly, in Indonesia case, the UN scores are used as part of admission into high school level. As a result, to get low scores would affect their academic records, admission of higher education level and teachers' or parents' expectations. In line with this, Ferman (2004) said that the fear of low scores is one of factors causing washback effects of the test on students' learning.

Regarding learning English communication, English communication skill was one of the tested skills in the O-NET and UN so that both ninth-grade Thai and Indonesian students focused on it. The results were similar to, for example, Ferman (2004), Akpinar and Cakildere (2013), and Lunrasri and Gajaseni (2014). Those studies found that the teacher emphasized their instruction and students stressed their learning on what the test tested. However, ninth-grade Thai students learned to communicate based on English textbooks while Indonesian learned from other materials too. This was a positive impact of the O-NET and UN because being able to speak English is one of the goals of English learning. Therefore, the students in both cases were promoted and encouraged to learn. However, in Thailand case, the teachers always used L1 in the classroom rather than English. In this regard, ninth-grade Thai students requested more English exposure from their teachers. In contrast, ninth-grade Indonesian students learned from various learning materials. Also, English was always used in the classrooms while L1 was used when complex explanations were needed. It illustrated that between ninth-grade Thai and Indonesian students, in average, the level of English proficiency and communicative competence of Indonesian student may be higher than Thai students. As a result, while they shared some factors and constraints of English learning, ninth-grade Thai students may have some problems related to English instructions from schools.

## CONCLUSION

To conclude, washback of a test is worth to study since it inevitably occurs and affects teachers and students in a myriad of ways. This present study not only showed the washback of the O-NET and UN on classroom settings, but it also revealed 1) what ninth-grade Thai and Indonesian students perceived about the tests, and 2) the similarities and differences between Thai and Indonesian contexts. Whereby, the results are useful in developing English teaching and learning at grade 9 level in both countries. In addition, it can be used to anticipate harmful washback from such tests.

Besides, there are few studies, which deal with the washback effects of national exams in Thailand and Indonesia. As a result, future studies should be conducted to investigate how to deal with such tests and

how to exploit the washback of the tests to improve English teaching and learning.

## REFERENCES

- Aydin, S. (2009). Test anxiety among foreign language learners: A review of literature. *Journal of Language and Linguistic Studies, 5*, 127-137.
- Akpinar, K. D., & Cakildere, B. (2013). Washback effects of high- language tests of Turkey (and ÜDS) on productive receptive skills of personnel. *Journal of Language Linguistics Studies, 9*(2), 81-94.
- Alderson, J. C., & Wall, D. (1993). Washback exist? *Applied Linguistics, 14*, 115-129.
- Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice*. Oxford: Oxford University Press.
- Badan Standar Nasional Pendidikan: BSNP. (2009). *Laporan BSNP Tahun 2009*. Retrieved from <http://bsnp-indonesia.org/id/wp-content/uploads/2011/02/Laporan-BSNP-2009.pdf>
- Bailey, K. M. (1996). Working for washback: A review of the washback concept in language testing. *Language Testing, 13*(3), 257-279.
- Bailey, K. M. (1999). *Washback in language testing*. ETS Educational Testing Series. Princeton, New Jersey. U.S.A.
- Benjamin, A. S. & Pashler, H. (2015). The value of standardized testing. *Policy Insights from the Behavioral and Brain Sciences, 2*(1), pp. 13- 23. <http://dx.doi.org/10.1177/2372732215601116>
- Center for Applied Linguistics (CAL). (2016). *Understand assessment: A guide for foreign language educators*. Retrieved 2 April, 2017 from <http://www.cal.org/flad/tutorial/impact/5washbackinstruction.html>
- Cheng, L. (2000). *Washback or backwash: A review of the impact of testing on teaching and learning*. Retrieved from <http://files.eric.ed.gov/fulltext/ED442280.pdf>
- Cheng, L. (2003). Looking at the impact of a public examination change on secondary classroom teaching: A Hong Kong case study. *Journal of Classroom Interaction, 38*(1), 1-10.
- Cheng, L., & Curtis, A. (2004). Washback or Backwash: A Review of the Impact of Testing on Teaching and Learning. In L. Cheng; Y. Watanabe & A. Curtis (Eds.), *Washback in language testing: Research contexts and methods* (p. 3-17). Mahwah, NJ: Lawrence Erlbaum Associates.
- Creswell, J. (2009). *Research design: Quantitative, qualitative, and mixed method approaches, edition*. Thousand Oaks, CA: Sage Publication.
- Ferman, I. (2004). The washback of an EFL National Oral Matriculation Test to teaching and learning. In L. Cheng, Y. Watanabe, & A. Curtis (eds), *Washback in language testing: Research contexts*



- and methods* (pp.191-210). Mahwah, NJ: Lawrence Erlbaum Associates.
- Frederiksen, J. R., & Collins, A. (1989). A systems approach to educational testing. *Educational Researcher, 18*(9), 27-32.
- Hamied, F.A. (2010). EFL assessment in Indonesia; national exam and quality of education. In n Moon, Y. & Spolsky, B. (Eds.). *Language assessment in asia: Local, regional or global?*(pp. 99-120). Seoul: Asia TEFL
- Kennedy, S., & Lui, R. (2013). Washback of a high-stakes English test in China: Student and teacher perceptions. *Concordia Working Papers in Applied Linguistics, 4*(10), 23-29.
- Lunrasri, Y. & Gajasen, C. (2014). Washback Effects of the Ordinary National Educational Test on English Language Learning as Perceived by Grade 9 Students. *Online Journal of Education, 9*(4), 226- 240.
- Messick, S. (1989). Validity. In: R. Linn (Ed.) (3rd Ed.) *Educational Measurement* (pp. 13-103). New York, NY: Macmillan.
- Messick, S. (1992, April). The interplay between evidence and consequences in the validation of performance assessments. Paper presented at the *Annual Meeting of the National Council on Measurement in Education*, San Francisco.
- Messick, S. (1994). The interplay of evidence and consequences in the validation of performance assessments. *Educational Researcher, 23*, 13-23.
- Messick, S. (1996). Validity and washback in language testing. *Language Testing, 13*(3), 241-256.
- Ministry of Education. (2009). *Basic Education Core Curriculum B.E. 2551(A.D.2008)*. Retrieved 3 April 2017 from <http://academic.obec.go.th/web/doc/d/147>
- National Institute of Educational Testing Service (NIETS). (2012). *The Operation Involves the Use of O- NET as One of Criteria in Determining the Learning Outcomes of Students*. Retrieved April 3, 2017 from <http://academic.obec.go.th/web/doc/d/776>
- National Institute of Educational Testing Service (NIETS). (2015). *O-NET (Ordinary National Educational Test)*. Retrieved 29 March 2017 from <http://www.niets.or.th/en/catalog/view/2211>.
- Pukmai, B. (2009). *The Relationship between EFL high school Learners' Scores on Teacher-made Test and the Ordinary National Educational Testing (O-NET)* (Unpublished M.A. thesis). Mahidol University, Thailand. Retrieved from <http://tdc.thailis.or.th/>
- Popham, W. J. (1993). Measurement-driven instruction as a quick-fix reform strategy. *Measurement and Evaluation in Counseling and Development, 26*(1), 31-34.
- Saukah, A. (2015). National Exam in Indonesia and Its Washback Effects. In F. A. Hamied, I. B. P. Yadnya & I. G. A. G. Sosiawati (Eds.). *Developing indigenous models of english language teaching and assessment* (p. 143-160). Bali: English Department Faculty of Letters and Culture and Post Graduate Study Program, Udayana University.
- Shepard, L. A. (1990). Inflated test score gains: Is the problem old norms or teaching the test? *Educational Measurement: Issues and Practice, 9*, 15-22.
- Shepard, L. A. (1991). Psychometricians' beliefs about learning. *Educational Researcher, 20*(6) 2-16.
- Shepard, L. A. (1992). What policy makers who mandate tests should know about the new psychology of intellectual ability and learning. In: B. R. Gifford and M. C. O'Connor, *Changing assessments: Alternative views of aptitude, achievement and instruction*. London: Kluwer Academic Publishers, pp. 301-27.
- Shepard, L. A. (1993). The place of testing reform in educational reform: A reply to Cizek. *Educational Researcher, 22*, 4, 10-14.
- Shih, C.-M. (2007). A new washback model of students' learning. *Canadian Modern Language Review, 64*(1), 135-161.
- Stake, R. E. (2006). *Multiple case study analysis*. New York, NY: The Guilford Press.
- Sukyadi, D., & Mardiani, R. (2011). The washback effect of the English National Examination (ENE) on English teachers' classroom teaching and students' learning. *K@Ta, 13*(1), 96-111. doi:10.9744/kata.13.1.96-111
- Sulistyo, G. (2009). English as a measurement Standard in the National Examination: some Grassroots' Voice. *TEFLIN Journal, 20*(1), 1-24. URL: <http://dx.doi.org/10.15639/teflinjournal.v20i1/1-24>
- Sullivan, L. E. (2009). High-stakes test. In *The SAGE glossary of the social and behavioral sciences* (Vol. 3, pp. 235-235). Thousand Oaks, CA: SAGE Publications Ltd. doi: 10.4135/9781412972024.n1189
- Taylor, L. (2005). Washback and impact. *ELT Journal: English Language Teaching Journal, 59*(2), 154-155. doi:10.1093/eltj/cci030
- The Greenwood Dictionary of Education 2<sup>nd</sup> edition*. (2012). Greenwood, California. U.S.A.
- UNESCO. (2010). Secondary education regional information base: Country profile. In *UNESCO Asia and Pacific Regional Bureau for Education*. Bangkok, Thailand: UNESCO.
- Virgo, E. (2016). An analysis of the impact of the English National Examination (ENE) on the teachers' teaching. *Satya Wacana Conference & Seminar, The 3rd ICETEACH*. Retrieved from <http://callforpapers.uksw.edu/index.php/iceteach/2016/paper/viewFile/101/100>
- Wall, D. (2000). The impact of high-stakes testing on teaching and learning: Can this be predicted or controlled? *System, 28*(4), 499-509

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

Zubaidi, N. (2014). *Ideal vs practice of Indonesian senior high school National Exam of English subject: Students' perception*. Retrieved from [http://www.academia.edu/7463607/Ideal\\_vs\\_Practice\\_of\\_Indonesian\\_Senior\\_High\\_School\\_National\\_Exam\\_of\\_English\\_Subject\\_Students\\_Perception](http://www.academia.edu/7463607/Ideal_vs_Practice_of_Indonesian_Senior_High_School_National_Exam_of_English_Subject_Students_Perception).