

The Washback Effect of the English National Examination (ENE) on English Teachers' Classroom Teaching and Students' Learning

Didi Sukyadi

English Education Department, Faculty of Language and Arts Education, Indonesia
University of Education, Dr. Setiabudhi 229, Bandung 40154, Indonesia
e-mail: dsukyadi@upi.edu; didisukyadi@live.com

Ridha Mardiani

Pasundan School of Teacher Training and Education, Cimahi
Permana No 32 B, Kota Cimahi 40152, Indonesia
e-mail: ridha.mardiani53@gmail.com

Abstract: Since 2005, Indonesia has administered the National Examination (NE), from which the result is used as a basis to decide the students' exit from secondary schools. As a high stake testing, the NE has caused various washback effects, both positive and negative. Consequently, the existence of the NE has been accompanied by heated debates and controversies. The pros believe that the NE may improve secondary education standards, while the cons argue that the NE will create injustice and unnecessary anxiety among the students. However, both the pros and cons base their arguments on common sense not supported by a sound reasoning. The present study tries to explore the washback effects of the English National Examination (ENE) in Indonesian secondary education context, involving three Secondary Schools categorized based on their NE achievement. The results of the study indicate that English teachers and students from the schools involved have different perceptions on the ENE. The English National Examination has an influential impact on teachers' teaching in the aspect of: activity/time arrangement, teaching materials, teaching contents, teaching methods, teaching strategies, ways of assessing, and on the feelings and attitudes of the students. The ENE also affects the students' learning in the classroom in which teachers mainly teach to test, practice the test and develop test-taking strategies. The dimensions of the washback effect of the ENE on both English teachers and students are negative and positive, strong, specific and for a short period time.

Key words: washback, English National Examination, curriculum, teaching materials, teaching methods, classroom teaching, practice the test, teaching to test, test taking strategies

Based on experiences, the implementation of the National Examination (NE) in one country will cause some impacts on its participants. In testing terminology, the impacts are recognized as washback effects. Washback is seen as the influence of testing on teaching and learning (e.g. Hughes, 1989; Alderson & Wall, 1993; McNamara, 2000; Brown, 2004; Gates, 1995; Cheng & Curtis, 2004; Cheng, 2005; Fulcher & Davidson, 2007). In addition, it can be defined as the effects of language tests on micro-level of language teaching and learning, i.e. inside the classroom (Bachman & Palmer, 1996; Bailey, 1996; Hamp-Lyons as cited in Hawkey, 2006; McNamara, 2000). Brown (2005) highlights that “washback is the degree to which a test affects the curriculum that is related to it” (p. 242). Shohamy (as cited in Bailey, 1999) summarizes four key definitions that are useful in understanding the washback concept: 1) Washback effect refers to the impact that tests have on teaching and learning; 2) Measurement driven instruction refers to the notion that tests should drive learning; 3) Curriculum alignment focuses on the connection between testing and the teaching syllabus; and 4) Systemic validity implies the integration of tests into the educational system and the need to demonstrate that the introduction of test can improve learning. Tests can also have effect beyond the classroom. The wider effects of tests on the community as a whole, including the school, is referred to as test impact (McNamara, 2000).

Watanabe as cited in Cheng & Watanabe (2004) conceptualizes washback as having the following dimensions:

- a) Specificity: washback may be general or specific. General washback means a type of effect that may be produced by any test; specific washback refers to a type of washback that relates to only specific aspect of a test or one specific test type.
- b) Intensity: washback may be strong or weak. If the test has a strong effect, then it will determine everything that happens in the classroom, and will lead all teachers in the same way toward exams. On the other hand, if a test has a weak effect, then it will affect only a part of the classroom events, or only some teachers and students, but not others.
- c) Length: the influence of exams, if it is found to exist, may last for a short period of time, or for a long time.
- d) Intentionality: Messick as cited in Cheng & Watanabe (2004) implies that there is unintended as well as intended washback when he wrote that judging validity in terms of whether a test does the job it is

employed to do...requires evaluation of the intended or unintended social consequences of test interpretation and use.

- e) Value or direction: examination washback may be positive or negative (Alderson & Wall, 1993; Hughes, 2003).

Tests that have important consequences will have washback (Alderson & Wall, 1993). In other words, tests will arguably have a greater influence on teaching and learning in a 'high-stakes' situation – i.e. one in which the test is typically used “to compare and rank individuals, schools or national systems” (Chapman & Snyder Jr. as cited in Cheng & Watanabe, 2004). A high-stake testing refers to test whose results are seen-rightly or wrongly- by students, teachers, administrators, parents or general public as the basis upon which important decisions are made that immediately and directly affect the students (Cheng, 2005). Greaney & Kellaghan (1996, p. 37) state, “An examination or test is said to have high-stakes attached to it when sanctions or rewards are linked directly to performance”. In addition, McNamara (2000) defines a high stakes test as “a test which provides information on the basis of which significant decisions are made about the candidates” (p. 48).

According to Cheng and Watanabe (2004) there is convincing evidence to suggest that examinations, especially high-stakes tests, have powerful washback effects on teaching and learning within different educational contexts. Pressure to improve students' test scores cause some teachers to neglect material that examination does not include and that teachers have negative feelings about standardized testing and narrowing of the curriculum (Smith as cited in Cheng & Watanabe, 2004). High stakes test mostly results in the phenomena of teaching the test and teaching to test. The former involves having some idea of what test questions will be and teaching the answers to the students. This method does not promote learning, and it is downright dishonest and could be considered as cheating. The later involves understanding which concept will be assessed and teaching those concepts to students.

The amount and type of washback depend on the extent to which: the status of the test, the degree to which the test counters to current teaching practices, what teachers and textbook writers are willing and able to innovate (Alderson & Hamp-Lyons as cited in Cheng & Watanabe, 2004). High stake test is usually used for more important decisions such as admissions, promotion, placement, or graduation decisions that directly

depend on test scores. For schools, high stakes decisions might include potential cuts in programs and funding that may also be directly dependent on score averages. The washback effect is obviously much stronger in high stakes situations than in low stakes situations. There are two types of washback which are well-known as the washback direction: negative and positive.

Negative washback is said to occur when test contents or format was based on narrow definition of language ability, and it constrains the teaching/learning contexts. It can also be said that negative washback is negative or undesirable effect on teaching and learning of a particular test, which means a poor test in which something that the teacher or learner does not wish to teach or learn and a mismatch between the content (e.g., the material/abilities being taught) and the test (Alderson & Wall, 1993; Brown, 2004). Washback becomes negative washback when there is a mismatch between construct definition and the test, or between content (e.g. the materials/abilities being taught). In line with this, Alderson and Wall (as cited in Cheng & Watanabe, 2004) note that test could be powerful determiners, both positively and negatively, of what happens in the classroom.

Some researchers strongly believe that it is feasible and desirable to bring about beneficial change in language teaching by changing examinations, which refers to so-called 'positive washback'. This term refers to tests or examinations that influence teaching and learning beneficially (Alderson & Wall, 1993); when testing procedure encourages 'good' teaching practices (Taylor as cited in Cheng, 2005). In this sense, teachers and learners have a positive attitude towards the test and work willingly towards its objectives. According to Pearson (as cited in Cheng & Watanabe, 2004), a test's washback effect will be negative if it fails to reflect the learning principles and course objectives to which the test supposedly relates, and it will be positive if the effects are beneficial and encourage the whole range of desired changes. Alderson and Wall (1993) stress that the quality of the washback effect might be independent of the quality of the test: any test, good or bad, may result in beneficial or detrimental washback effects.

Two studies conducted by Pizarro (2009) and Spratt (2005) show in detail what areas in teaching and learning affected by washback, namely curriculum, materials, teaching methods, feelings and attitudes, learning, teaching strategies and teaching contents.

- 1) Curriculum: the results of some studies carried out on washback showed that examination has had a demonstrable effect on the context of language lesson-narrowing of the curriculum to those areas most likely to be tested (Alderson & Wall, 1993).
- 2) Teaching materials: material in this case is the exam-related textbooks and past papers. As the exam getting closer there is greater use of past paper and commercial exam-public examinations (Alderson & Wall, 1993).
- 3) Teaching methods: in this study refers to teaching approaches and techniques. Alderson & Wall (1993) stated that a study in Sri Lanka showed that the exam had virtually no impact on the way that teachers teach. Teaching toward the high-stakes EFL exam led the teachers to teach through simulating the exam tasks or through carrying out other activities directly aiming at developing exam skills or strategies (Shohamy, 1996).

The National Examination (abbreviated into NE) is the latest form of a school leaving examination in Indonesia starting from 2005 until now. It can be defined as a test to measure and evaluate the students' competence nationally by the central government after the process of teaching and learning (The Regulation of the Minister of Education 2005, p.1). The NE is implemented as a way of improving national education quality, which is supported by Government Regulation Number 19 year 2005 on the National Standards of Education. Since the school year 2003/2004 the English National Examination in Indonesia has included listening and reading skills with the number of the items of 20 listening questions and 40 reading questions. The time allotted to do the exam was 120 minutes. In the school year 2008/2009, ENE test items consisted of 15 listening questions and 35 reading questions. As pros and cons concerning the National Examination are getting more intense, the present study tries to explore the washback effects caused by the ENE in Indonesia, including whether the washback is positive or negative, the areas of washback, and its dimensions.

METHODS

This study used a qualitative method, using multi sites contexts combined with several data collection techniques including observation,

interviews, questionnaire and documents. Three Secondary Schools were chosen and categorized into a) high-level achiever b) moderate- level achiever and c) low-level achiever. The categorization of schools was based on NE scores for the last four years. Participants in the study were: (a) the twelfth grade English teachers from three Secondary Schools as the main group; (b) the twelfth grade students from each school chosen purposively for the observations; (c) the tenth and the eleventh grades English teachers to obtain supplementary data; and (d) the tenth and the eleventh grade students from each school as the secondary groups; and (e) school principals. The researcher applied inductive analytical approach by searching for emerging patterns within and across the twelfth English teachers’ classroom teaching, the tenth and the eleventh grade English teachers’ classroom teaching.

FINDINGS AND DISCUSSION

Teachers’ Perception towards English National Examination

Data for English teachers’ perceptions on the ENE were obtained mainly from teachers’ questionnaires and interviews. The following findings, derived from teachers’ questionnaires, were developed in the form of open-ended questions. Evidently, these findings confirmed that the tenth, the eleventh and the twelfth grade English teachers’ perceptions of the English National Examination vary. There are two categories of perception, positive and negative as can be observed in the following tables 1.

Table 1. English Teachers’ Positive Perceptions on NE

Category	School A			School B			School C
	R1	R2	R3	R1	R2	R3	R1
National Standard	v			v		v	v
School Rank		v			v		
School Quality		v	v				
School progress			v				
Students’ Achievement					v		
Forcing students to study		v		v	v		v
Cognitive Skills						v	

Table 2. English Teachers’ Negative Perceptions on NE

Category	School A			School B			School C
	R1	R2	R3	R1	R2	R3	R1
Quality improvement	v			v			v
Dishonesty		v					
The only standard			v		v	v	v
Unfairness		v					v
Useless				v			v

Table 1 shows the data from teachers’ questionnaires, which stated positive perceptions of NE implementation. Four respondents perceive NE as important for national standards. Two respondents stated that NE served to examine school rank, quality, progress and students’ achievement. One positive point in their perception is that NE compels the students to study more to improve their academic knowledge and skills, vocabulary, and reading comprehension. However, we see that school progress and students’ achievement are supported by two respondents out of seven, indicating that the respondents do not put emphasis on the measurement of their learning achievement. A more interesting finding is revealed from Table 2. It is shown in Table 2 that there are three respondents who stated their objections that NE is the only standard for students to pass school leaving examination. They also raised objections about the fairness of NE that the length of the students’ studying for three years is only determined in five days of NE implementation and finally one respondent held that the result of NE is futile since the students have to take another test to enter colleges.

Areas Affected by English National Examination

In this study, the areas affected by the NE on English teachers’ teaching are set up and limited to the areas that have previously been investigated by some researchers on washback studies. It can be identified that there are seven areas of English language teachers’ teaching affected by the public or the national examination: 1) Activity/Time arrangement; 2) Teaching materials; 3) Teaching contents; 4) Teaching methods; 5) Teaching strategies; 6) Ways of assessing; and 7) Feelings & attitudes.

Table 3. The Areas of the Twelfth Grade English Teachers' Classroom Teaching Affected by the Washback Effect of ENE

Areas of Teaching	School A	School B	School C
Activity/time arrangement	Additional time for students whose scores below passing grade	Additional time for ENE preparation	Additional time for extra class hours
Teaching Materials	More reading exercises, textbooks, text types (genre), supplementary materials, ENE past papers, commercial books for ENE preparation	More text types, commercial books for ENE preparation, reading texts, ENE past papers	More textbooks, supplementary materials from the Internet and newspapers
Teaching method	Explanation, exercises, listening practice, test practice, ENE try out.	Practicing the test, exercises, questions and answers, teaching to the test from ENE past papers	Story telling, watching movies, test practice, listening practice
Teaching content	Teaching objectives, exit competencies	Curriculum, text types	curriculum
Teaching strategies	Exercise more on ENE format	Exercise more on ENE format, test taking strategies	Exercise more on ENE format, developing communicative competence
Way of assessment	Mid semester, pre-final exam, final exam, ENE try out, Pre ENE	Mid semester, pre-final exam, final exam, ENE try out, Pre ENE	Mid semester, pre-final exam, final exam, ENE try out, Pre ENE
Feeling and attitudes	Nervous, stressful, pressure to score 9 for English,	No pressure	Pressure from the students

Table 4. The Areas of the Tenth and Eleventh Grade English Teachers' Classroom Teaching Affected by the Washback Effect of ENE

Areas of Teaching	School A	School B	School C
Activity/time arrangement	No additional/extra time	No additional/extra time	No additional/extra time
Teaching Materials	Textbooks	Textbooks	Textbooks
Teaching method	Various methods	Dialogues	Various methods
Teaching content	Curriculum, supplementary materials	Curriculum, life experiences	Curriculum, supplementary materials
Teaching strategies	Developing communicative competence	Group discussion	Developing communicative competence
Way of assessment	Mid semester, pre final exam, final exam, giving points or scores	Mid semester, pre final exam, final exam	Mid semester, pre final exam, final exam
Feeling and attitudes	No pressure	No pressure	No pressure

In table 3, data obtained from teachers' questionnaires, interview and classroom observation verify the presence of the washback effect of ENE on the XII grade English teachers' classroom teaching. Firstly, the washback effect in the area of activity or time arrangement. All the XII grade-English teachers in all three secondary schools altered their activities and class arrangements in response to ENE and gave additional or extra time in their teaching. They skipped language classes to prepare for the test instead (Bailey, 1996); allocating time to the skills needed in the test (Pizarro, 2009). The second area affected by washback is teaching materials. The teachers replaced class textbooks with worksheet identical to previous years' tests. From the sample of five sessions of classroom observation, it was found that the main activity in the classroom was "teaching to the test" or "practicing the test". The teachers tried out ENE materials, discussed ENE previous tests and particularly drilled listening skills which would be tested in ENE.

In facing the ENE, all respondents changed their teaching method into ENE preparation in which classroom activities were mainly focused on practicing the test. It shows that the washback effect of the National Examination on teachers' teaching method is negative, because teachers will do whatever methodology they feel most expedient to help them to prepare their students for the English National Examination (Alderson & Wall, 1993). The respondents stated that as a strategy, they recommended the students to do more exercise in the ENE test format to improve their reading skills. In this case, teaching to test and testing to test occurred. Students were practicing "test-like" item similar in format to those on the test (Bailey, 1996; Shohamy *et al.*, 1996). Those materials were acquired from previous ENE materials available on the web, or those commercially prepared by many authors.

All respondents assessed their teaching by giving mid semester test, pre final exam and final exams. The data from teachers' questionnaires show that the basic function of ENE Try-Outs is to know their current performance in regard to ENE passing grade so that it can motivate the students to learn and work harder. Speaking and writing skills were tested in school through performance based assessment. The tests did not play any significant part in deciding whether a student could leave the schools or not. Therefore, the washback effects of tests administered by schools were not quite significant. On the other hand, high-stake testing such as NE indeed impacts the feeling of pressure. Referring to the study conducted in North Carolina by Jones, M., Jones, B., & Hargrove (1999), the finding confirms that a high-stakes test compels the teachers and students, particularly from schools in the category of high-level achiever, the schools and teachers intend to maintain or exceed previous test achievements. These findings are contradictory to the data in Table 4, which show that the areas of English teachers' classroom teaching in the tenth and the eleventh grade were not affected by the washback effect of the ENE.

Most of the students from school B and school C complained that the ENE had made them stressful. This is a common reality because students would encounter stress anyway in their educational lives. It is quite an uncommon phenomenon when the students from school A state that ENE is not a burden for them; ENE does not make them anxious or worry. They do not consider ENE as a nuisance, but motivates them to learn more diligently. In this case there is a positive washback effect of ENE on the students' learning.

Some aspects of the students' learning in School A, B, and C share the same notion, such as: the students were seriously preparing for ENE since the 2nd semester of the twelfth grade; ENE had affected the material taught by the teachers particularly when they had ENE try out or pre-ENE test; ENE has taken a great portion in students' learning; the students identify that teachers have taught them according to the Graduate Competence Standard; and ENE preparation would make the students ready to face ENE.

The data from students' interviews also showed the same notion on the washback effect of ENE on students' learning in school A, B and C. The respondents from school A stated that the students practiced listening skills in language laboratory indicating that they had a lot of exercises on ENE particularly in the 2nd semester when ENE administration due date was getting closer. Respondents from school B described that the students learned and did a lot of exercise on reading comprehension and text types, in which grammar rules were only learnt in text context. Respondents from school B and C agreed that their students practiced listening skills intensively in the 2nd semester of the XII grade. Sometimes the teachers asked them to have note-taking, explained and practiced a lot of tests. When they learned speaking, they practiced reading and memorizing dialogues, while for writing they learned and practiced writing paragraph or made summary of stories or texts. Students from school C indicated that they learned English proportionally; then, they focused all their efforts on ENE material in the 2nd semester of the XII grade, particularly for practicing the test.

The Dimensions of the Washback Effect

The basis in data analysis to reveal the dimension of washback is the framework adopted from Watanabe (as cited in Cheng & Watanabe, 2004) with the following findings:

- a) The specificity of the washback effect of ENE existing in the context of this study is specific to one type of washback, or one specific test type. In school A, R#1 had spent more time instructing reading than communication skills, because he thinks reading takes greater part to be tested on ENE and he also emphasized vocabulary teaching. In school B, R#1 emphasized her materials on reading genres starting from its characteristics, purposes and text types, and she skipped certain material

not tested on ENE. In school C, R#1 gave a lot of exercises on listening skills and reading texts.

- b) The intensity of the washback is strong, because ENE has a strong effect and determines everything happened in the classroom, and lead all teachers to teach in the same way toward the examination.

The data from teachers' questionnaires, interviews and classroom observation confirm the strength of the washback effect. Because ENE could be regarded as a high stakes test, it has affected strongly of what has happening in the teachers' classroom teaching.

- c) The length of the washback lasts for a short period of time.

As it can be seen from the data of teachers' interview and classroom observation, it is noted that the influence of ENE started since the 2nd semester in the twelfth grade which means that it existed only while the students were preparing for ENE and the influence would disappear after ENE had been administered and the students had passed the exam.

- d) The intentionality of the washback occurring in this study is intended as well as unintended washback.

Based on the data from the interview and classroom observation, ENE has caused intended washback in the case that the subject matters tested in ENE, including English, cover the material delineated in the curriculum. Data from classroom observation reveal that unintended washback occur when teachers' focus on "teaching to the test" became the objectives of the teaching. This happened in the 2nd semester, particularly when the due date of ENE administration were getting closer.

- e) The attitude towards washback existing in this study is both positive and negative.

There is a kind of negative washback effect of ENE on the twelfth grade English teachers' teaching materials. Teachers from school A, B, and C agreed on the fact that they had instructed reading comprehension and ENE past papers. The washback effect of ENE on the twelfth grade English teachers' teaching methods is negative as teachers would do whatever methodology they felt most expedient to help them to prepare their students for ENE (Alderson & Wall, 1993).

After analyzing data from students' questionnaires and interviews, there are some similar findings found in the context of study of school A, B, and C. Below are the findings from data analysis for school A, B and C.

- The washback effect of the National Examination strongly exists in students' learning.
- There is a negative washback effect of the National Examination on students' learning particularly for learning material in which they mainly prepare to face the National Examination.

The specificity of the washback effect existing in the context of this study is specific to a type of washback related to only one specific aspect of a test or one specific test type. The students learn reading genres including text types and vocabulary and a lot of exercises on listening skills and reading texts.

- The intensity of the washback effect of the ENE on students' learning was strong.

The intensity of the washback effect of the ENE on students learning was strong, because the ENE had a strong effect and determined everything happening in the classroom. This is caused by the stakes of ENE as a high-stake test which leads the students to learn toward the examination.

The intentionality of the washback occurring in this study is intended as well as unintended washback. When the data from the interview and classroom observations were scrutinized closely, it can be seen that the ENE had caused intended washback in the case that all subject matters tested in the ENE, including English, cover the material delineated in the curriculum. The data from classroom observations reveal that unintended washback had also occurred when students' "learn the test" or "practice the test" becomes the objectives of the teaching. This happens in the 2nd semester, particularly when the due date of ENE administration was approaching.

- There is a kind of negative washback effect of ENE on the twelfth grade students' learning.

The data from classroom observations show that the main activity in the classroom was "teaching to the test" (Alderson & Hamp-Lyons, 1996), meaning that teachers use materials that mimic the format and cover the same curricular territory as the test (Smith, 1991) and "practice the test" in which the students practice on items of the test itself; the teacher replaces textbooks with worksheet identical to previous years' tests (Shohamy et al., 1996).

CONCLUSIONS

English teachers involved in this study have positive and negative perceptions on the ENE. Overall they understand that the ENE is important for national standard, for measuring the schools' progress, improving the quality of education, and determining students' graduation. The areas affected by the ENE are: 1) Activity/Time arrangement, 2) Teaching materials, 3) Teaching methods, 4) Teaching contents, 5) Teaching strategies, 6) Ways of assessing, and 7) Feelings and attitudes. The washback of the ENE affect the areas of the twelfth grade English teachers' classroom teaching, but not for the tenth and the eleventh grades. This indicates that the washback effect of the ENE only takes place when the students and teachers feel that the ENE preparation is a must. However, the school of high-level achievers tends to have more positive washback in the teachers' teaching and students' learning, while the school of moderate-level achievers tends to have more negative washback particularly in teaching materials and teaching contents. Worse, the school of low-level achievers tends to have more negative washback than their higher achiever cohorts. This is due to the students' condition, which leads the teacher to teach in accordance with their need and capability. If they do not try hard, teachers from low achiever school are worried to have more students failing the national exam.

REFERENCES

- Alderson, C. J., & Wall, D. (1993). Does washback exist? *Applied Linguistics*, 14(2), 115-129.
- Alderson, C. J., et al. (1995). *Language test construction and evaluation*. Cambridge: Cambridge University Press.
- Alderson, C., & Hamp-Lyons, L. (1996). TOEFL preparation courses: A study of washback. *Language Testing*, 13(3), 280-297.
- Bachman, L., & Palmer, A. (1996). *Language testing in practice*. Oxford: Oxford University Press.
- Bachman, L. F. (2004). *Statistical analyses for language assessment*. United Kingdom: Cambridge University Press.
- Bailey, K. M. (1996). *Language teacher supervision*. New York: Cambridge University Press.

- Bailey, K. M. (1999). *Washback in language testing*. Retrieved May 2, 2008, from <http://www.toefl.org>
- Brown, D. H. (2004). *Language assessment principles and classroom practices*. Longman: Pearson Education
- Brown, J. D. (2005). *Testing in language programs*. New York: McGraw-Hill Companies.
- 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* (pp. 3-17). Mahwah, New Jersey: Lawrence Erlbaum.
- Cheng, L., & Watanabe, Y. (2004). *Washback in language testing research contexts and methods*. New Jersey: Lawrence Erlbaum Associates.
- Cheng, L. (2005). *Changing language teaching through language testing*. Cambridge: Cambridge University Press.
- Fulcher, G., & Davidson, F. (2007). *Language testing and assessment: An advanced resource book*. Canada: Routledge Applied Linguistics.
- Gates, S. (1995). *Exploiting washback from standardized tests*. In J. D. Brown, & S. O. Yamashita (Eds.), *Language testing in Japan* (pp. 101-106). Tokyo: Japanese Association for Language Teaching.
- Greaney, V., & Kellaghan, T. (1996). *Monitoring the learning outcomes of educational systems*. Washing D. C.: The World Bank.
- Hawkey, R. (2006). *Impact theory and practice studies of the IELTS test and Progetto Lingue 2000*. Cambridge: Cambridge University Press.
- Hughes, A. (1989). *Testing for language teachers*. Cambridge: Cambridge University Press.
- Hughes, A. (2003). *Testing for language teachers* (2nd ed.). Cambridge: Cambridge University Press.
- Jones, M. G., Jones, B., & Hargrove, T. (1999). *The unintended consequences of high stakes testing*. Boulder: Rowman and Littlefield.
- McNamara, T. (2000). *Language testing*. Oxford: Oxford University Press.
- Pizarro, M. A. (2009). Does the English teaching in the Spanish university entrance examination influence the teaching of English? *English Studies*, 90(5), 582-598.
- Shohamy, E. (1996). Test impact revisited: Washback effect over time. *Language Testing*, 13, 298-317.

- Smith, M. L. (1991). Put to the test: The effects of external testing on teachers. *Educational Researches*, 20(5), 8-11.
- Spratt, M. (2005). Washback and the Classroom: The Implications for teaching and learning studies of washback from exam. *Language Teaching Research*, 9(1), 5-29.
- Wall, D. (1993). *The impact of national test on teaching: A cases study*. Unpublished manuscript, Institute for English Language Education Lancaster University.