

The Impacts of *not having tests* on Teachers and Students: More Influential Factors Mediating Washback Being Processed than Teachers?

Byung Rok Lee
(Seoul National University)

Lee, Byung Rok. 2018. The impacts of *not having tests* on teachers and students: more influential factors mediating washback being processed than teachers? *SNU Working Papers in English Linguistics and Language* 16, 83-100. Tests, especially high-stakes tests exert power on stakeholders' lives in every possible way, which might be either positive or negative, i.e. washback. Meanwhile, they are considered to have mainly negative impacts. Teachers teach to the tests by only focusing what is to be assessed; students do not allocate their time and effort to studying materials other than what will be on tests; a great deal of resource in the society will be invested solely to things related to tests. Granting it to be true, tests keeps motivating test-takers to upgrade knowledge and talent, giving opportunities to get a level playing fields for stakeholders. What if there were no *testing*? With respect to teachers, if it were not for *testing*, in what way would teachers be impacted in terms of classroom behaviors and their materials for class? Would there be any particular aspects of language that teachers were likely to put into action? Would they actually integrate them into their classroom activities? Then would *not testing* also have positive or negative impacts on students? If students stay motivated to keep pursuing their goal even in the absence of *testing*, what would be the factor preventing them from being deviated off course? Even though it seems that it is *personal factors* like teachers that are associated with negative washback, would there be any other factor other than that? Future research may try to investigate other phenomena that might take place but for *testing*: if testing were to be excluded, then would test-takers less likely depend on memorization? (Seoul National University)

Keywords: washback, negative, impacts, mediating factors, not testing, motivation

1. Introduction

The limitedness of resources and job openings force people to compete with each other and to come up with reasonable means to distinguish more appropriate human resources from the less. Students devote a great deal of time and effort to preparing for university admission, test-

takers to getting accepted by large-companies. Still, the society lets them continue to compete, finds better-suited persons, and places them into the economically-scarce positions. *Testing* plays an important role in probing the battle field and locating them into the right place. As the integration of almost every field has been commonplace in globe, some large-scale tests like Test of English as a Foreign Language (TOEFL) and International English Language Testing System (IELTS) have been given more and more important roles, and language testing is not alone. Testing is believed to be beneficial to keep students and test-takers motivated to improve their skills, talents and aptitudes, and makes it easier for corporations to find more appropriate persons to their enterprise. While *testing* itself evolves, so do test-takers and teachers so that they can have upper hands. We cannot emphasize the importance of *testing* too much and its impacts on students, their parents, teachers, material production field, and the society as a whole have developed. The concept of *washback* has been used to describe the impacts of *testing* on teaching and learning, admitting that there are some varieties of defining the term among researchers. Unfortunately, *washback* has been witnessed to influence the society mainly in a negative way. Students only focus their energy and time on what is to be assessed and tested by relying on memorization. They are sometimes blamed for only learning testing techniques and not learning the skills or language necessary. This situation has been interpreted to limit their creativity and critical thinking and appears never to stop. This is also the case with teachers. They *teach to the tests* so as to fulfil the societal requirements and narrow their curriculum offerings of instruction by ignoring the material that are considered to be less important in *testing* and reacting reluctantly to the introduction of new instructional practices. Policy-makers make full use of assessment and testing to exert their political power. They seem to know that if there happen any changes in *testing*, the society will react in a sensitive manner and it is to be persuaded to act in intended way.

There have been some movements to get rid of *testing* and give students and test-takers more opportunities to participate in group assignment and to improve their critical thinking and creativity. Korean government have recently introduced so-called *free semester system* to middle schools, in which students are exempted from *testing* and it gets substituted with assessments. Even though the impacts of *not testing* even for a while need longitudinal investigation to figure out whether the implementation of the policy brings about positive or negative results to the entire society, it is clear that *testing* has been stigmatized to exert negative influence on stakeholders as well as the society. Then would *not testing* be an answer to this situation? Would *not testing* have impact on teaching and learning in a similar way as *testing* has been?

2. Literature Review

No matter what field people are, they are subject to a certain type of *testing*. That may take a form of assessment, testing, or evaluation. The results they get from testing may decide which job they may have in the future, which university they will attend, or how financially affluent they will be. It is no wonder that they get to be sensitive about the results. In other words, testing has impacts on such stakeholders as students, their parents, test-takers, teachers, and so on. The impacts may affect the stakeholders in a positive or negative way, which are called *washback*. Some researchers tried to formulate the exact definition of washback. Messick (1996) defined *washback* as ‘the extent to which a test influences language teachers and learners to do things they would not necessarily otherwise do that promote or inhibit language learning.’ The existence of *testing* will change the way teachers behave and students learn in the classroom. Note that the general education is not the exception. Subjects like math, history, etc. may have stronger

influence on the society depending on the policy increasing or decreasing the importance of the tests for those subjects.

There have been some efforts to get a more specific definition about *washback*. Wall (1997) made an attempt to distinguish the test impact and test washback. The former indicates that tests can have impact on almost everything and everyone, including individuals, the schools, the education system, society, policies, activities in classrooms, test-takers, and so forth. Any change of tests will modify the way students learn, parents support and even the way material publishers print the books. The latter meaning the effects of tests on teaching and learning, admitting that there is quite a little overlap in between.

Considering the fact that test-takers are subject to testing itself, it is doubtless that tests impact teachers, classrooms, and students (Alderson & Wall, 1993). Students are most likely to choose to study what is deemed to be important for a certain testing, and in order to bring about great achievements, they will decide to teach what students are supposed to study according to the announcement of the government. McEwen (1995) even mentioned: 'what is assessed becomes what is valued, which becomes what is taught.' The fact of testing being considered to be important may lead to impacts not only on education, in general education and in language testing, but also other aspects like policies. In some countries, when the new administration takes office, they tend to change policies about education with some purposes, and it is assumed that finding the exception would be difficult. In addition, what determines what is taught, how it is taught, what is learned and how it is learned is testing rather than the officially stated curriculum (Madaus, 1988). When negative impacts of *testing* get noticed and it is noticed that what is being happening in the classrooms should not be the way it should be now, stakeholders who have rights to detect and modify the defects need to take into account the revision of testing itself. According to Alderson (1986), the persons concerned should innovate the language curriculum through the innovation in language

testing. It means that as long as the current language testing takes the form, teachers and students are not willing to change the way they study and teach. Elton and Laurillard (1979) also stressed that the quickest way to get the students to change the way they study is to change the assessment system. The influence of testing does not remain in the field of education: if policy-makers are to make people change their approaches to life, they can take advantage of testing and assessments. Madaus (1985) noted: testing and assessment is “the darling of policy-makers.”

Many researchers have focused on the fact that negative *washback* has been more witnessed than the positive one. In his research, Vernon (1956) argued that examinations distort the curriculum. Examinations exert such powerful influence that without changing the curriculum itself, they can actually change the way what is taught and how it is taught and what is learned and how it is learned. In order to get good grades, students most likely focus on practicing the exam technique, ignoring the importance of language learning activities (Wiseman, 1961). In the end, students may lose the ability to think in a critical way, only by making full use of memorization (Anderson et. al., 1990). When it comes to high-stakes testing, test-takers are under pressure that they cannot be afforded plenty of time for getting scores required for getting employed or being admitted by universities they want to attend. It is not just the students that are affected in a negative way by testing, but also teachers are vulnerable to the negative *washback*. When it comes to high-stakes testing, it affects teachers directly and negatively (Noble and Smith, 1994). High-stakes tests wields so much power that it influences personal fate as well as the future of the Empire (Spolsky, 1995). The important of getting student to achieve good results forces teachers to concentrate on testing programs, which considerably reduces the time available for instruction, and narrows the diversity of instruction. What is worse, that may reduce teachers’ ability to teach content and to use methods and materials that are incompatible with

standardized testing format (Smith, 1991). Chapman and Snyder (2000) pinpointed the teachers' tendencies to *teach to the test* are the factor inhibiting the introduction of new instructional practices.

Having said all the negative washback associated with testing, it should be noted that testing in itself is so indispensable and necessary a means to distinguish the well-prepared test-takers from the less prepared ones that it should be given enough credits that it deserves. Bray and Steward (1988) noted in their research that testing is an appropriate device to keep students motivated to develop their talent, to get schools and colleges to keep trying to improve their performance, and last but not least, to counter to some degree nepotism, favoritism, and even outright corruption in the allocation of scarce opportunities. According to Ralph Tyler (1960), "Tests are here to stay and their uses will increase. ... Because testing can increase motivation and provide an additional source of reinforcement to learning, and because it can help clarify the objectives of the school and focus effort on them, testing should be used as a positive factor in the educational program. ... Pupils and teachers... should have continuing experience with appropriate tests to gain confidence in their ability to learn and to perform and to utilize the reinforcement potential of tests as an important dynamic factor in learning (p.16)."

Messick (1988) argues that if tests were to be eliminated, the multiple purposes that the current tests serve will still survive and need to be addressed by other means. He lists in his research what would happen if it were not for tests. Schools would need to find alternative ways to fill the place held by tests, ending up with subjective appraisals. They would most likely be faced with issues regarding reliability, losing comparable yardsticks. Job training and education fields would not be able to assess students, losing their valuable instructional time to preliminary observation.

3. Research Questions

While tests are here to stay with its multiple advantages, they are criticized to have negative washback to their stakeholders, and when it comes to high-stakes tests, the criticism against them takes off. Considering their advantages and disadvantages, would not it be reasonable to address a question, “would teaching and learning be different if it were not for tests (Watanabe, 2004)?” Ebel (1996) also argued that it is important “to weight the potential social consequence of not testing at all.”

There have been movements in some countries to get rid of tests in the school semester by focusing on the negative impacts that tests possess: they deprive students of the opportunities to improve creativity and think critically. In Korea, the ministry of education implemented *free semester system* into the first semester for the first grader in middle schools as of 2016, which is expected to be switched to *free year system*. By saying the term *free*, the system indicates that first grader in middle schools are exempted from tests for the first semester, when instead of being under pressure of spending a great deal of time and effort preparing for tests, they are supposed to participate in social activities with a view to improving talent and to finding aptitude. Tests, which have been charged of their negative washback on the teaching and learning, have been replaced with assessments and evaluation for other means like group activities in the place of tests. Furthermore, the government in Korea has gone on to enlarge its plans to reinforce the so-called *revolutionary schools*, which are to minimize the influence of tests. Even though those large-scale attempts have been faced with harsh criticism and protests from stakeholders, it is notable that tests are actually faced with eradication in some places.

Cheng (1999) argued that teachers are the prominent factor mediating the process of *washback* being produced. There are other factors other than *personal factors* including teachers. First, test methods, test

contents, skills tested, purpose of the test, decisions that will be made based on test results, etc. are involved in the *washback*, i.e. test factors. Second, *personal factors* involve teachers' educational backgrounds, their beliefs about the best methods of teaching and learning, and so on. Teachers have tendency to remain unchanged by the introduction of new instructional practices and the older and less experiences teachers were, the more negatively they were likely to react (Fish, 1988). Third, the school setting in which test preparation is being carried out is considered to be the factor as well, *micro-context factor*. Fourth, *prestige factor* involves stakes of the test, status of the test within the entire educational system, etc. Last but not least, there is a *macro-context factor*, which indicates the society on the whole. A society places its own value on a certain thing like economic success or leisurely life style. The propensity may be strongly associated with *washback*.

The research will focus on what *not testing* brings about and how teachers and students reacts to the situation. Therefore, the research questions are:

1. Would *not testing* also have positive or negative impacts on students? If they do not change their approaches to studying even in the absence of *testing*, what would be the factor mediating the process of washback?
2. If it were not for *testing*, in what way would teachers be impacted in terms of classroom behaviors and their materials for class?
3. Would there be any particular aspects of language that teachers were likely to put into action, if they are capable of handling any change in language instruction and have a good command of English?

4. Methodology

In order to compare the impact of *testing* and *not testing*, the testing of concern should be high-stakes one. It needs to be thought of as exerting a profound influence on teaching and learning. In terms of language testing, Test of English as a Foreign Language (TOEFL) and International English Language Testing System (IELTS) may be appropriate tests for the research. And yet, in the absence of either of the testing or both, stakeholders are expected to switch to other means of testing due to the fact that there are alternative ways even if they may be considered to be not as reliable and effective as TOEFL and IETLS. It means that these large-scale language tests cannot be used in the research. The appropriate test should not have direct alternatives and a great deal of attention and economic resource should be devoted to its application.

Like any other countries, university admission is considered to be one, or the most important high-stakes testing in Korea. Not only do test-takers prepare for the college entrance exam in high schools, but they, in fact, start studying for the test as early as when they are in elementary schools. This test has been in place, exerting a good deal of power on Korean society for ages, while it being changed in forms. It is, of course, so important a means in Korea that it is impossible to go without it and it cannot be replaced with any other type of assessment and evaluation. However, the Korean ministry of Education has recently decided to get examination to be excluded from the first semester in the course of middle schools, which means that students are exempted from *testing* during the period and is called *free semester system* (FSS). The government has gone on to prolong the period from a semester to a whole year, which is the *free year system* (FYS). The purpose of those systems is to get the students out of the pressure of *testing* and to get rid of negative *washback* of *testing*. Those systems have been expected to give students opportunities to get involved in

social activities and to participate in group works and some other types of activities to improve talents and critical thinking. Even though *testing* itself has been replaced with other type of assessments and evaluations, they are considered to be less motivational and influential than *testing*. Similar type of endeavor has been witnessed by the Korean government's policy to reinforce the implementation of so-called *revolutionary school system* (RSS). FSS, FYS and RSS are believed to be the only, if any, possible environment that can be thought of not to have *testing* for a while, and that is still high-stakes *testing* and still does not have any specific alternatives. With the points mentioned above in mind, this research is to make use of university entrance exam in Korea as a target *testing*, the sections of which English language *testing* will be investigated.

4.1. Participants

4.1.1. Teachers.

Two groups of teachers will participate in the research. First group will consist of five Korean teachers who teach English and work in middle schools in Korea. They are anticipated to have experienced the advantages and disadvantages of FSS. They have utilized assessment and evaluations to measure educational achievement of students other than *testing* for a semester. Second group comprises two Korean teachers who teach English and work or worked in high schools in Korea and are deemed to have a good command of English. They are considered to be capable of integrating speaking or writing into their classroom curriculum if necessary. In order to fulfil the requirement, teachers who are also involved in the department of English Language and Literature at the same time will be chosen due to the fact that they have proven their English proficiency with more than 800 in scores from the Test of English Proficiency Test (TEPS) by Seoul National University.

4.1.2. Students

The research will involve two groups of students. First group will have 30 students who will be starting to study in high school in 2019. They have experienced FSS back in 2016, when they were freshmen in middle school. Compared to predecessors, they have one semester out of six, when they were exempted from *testing*, which were replaced with assessments. Second group does not actually participate in the research in person. The group will be all the students in Korea who were the freshmen in high school in 2018. They do not have to take part in the research in person, because their official results have been made public by the Seoul Metropolitan Office of Education on March, 2018.

4.2. Experiment

4.2.1. Teachers

The two groups of teachers will be required to fill in the questionnaires and to be interviewed. The questionnaire for the group of middle school teachers will ask them about what they have felt about advantages and disadvantages of FSS, and what they expect about FYS, which will include questions like: 1) what did they feel about not having tests, 2) how did students react to not having tests, 3) were the assessments and evaluation proper means to replace for testing, 4) did students keep motivated to study as hard as they would with *testing*, 5) are they positive or negative about the transition of FSS to FYS, 6) were parents get persuaded by the advantages of FSS and satisfied with their children doing what is required by the system and not doing the preparation study for future tests coming up, and so on. These questions are to get the opinions from stakeholders who have got involved personally in the system, so that their opinions are considered to hold more meaningful messages. The interview given to the group has the similar purposes, but still is to be expected not to be redundant in that this interview will

give them a chance to support their opinions shown on the questionnaires.

The second group of teachers who are working or have experienced working in high school are also to be given questionnaires and interview. They will be given questions like: 1) were it not for *testing* and if they were given freedom to include what they think would be more helpful as a means of education, what would they include in the curriculum, 2) if speaking and writing are required by students, would they be willing to accept them as a part of curriculum, 3) would they be able to reduce the reservations that involving those sections may discriminate against the poor, 4) instead of *testing*, what do they think is the most appropriate evaluation methods, 5) how good or bad would the students be at English who are exposed to FSS or FYS, 6) if they were willing to change in the absence of *testing*, but still students were unwilling and kept the traditional learning practices, what would be factor mediating the *washback* of testing being produced, and so on. After finishing the questionnaires, they will be interviewed to give them a chance to support their opinions. Compared to the former group, they are considered to be more involved with a high-stakes test, which is a university entrance examination in Korea.

Teachers are believed to be reluctant to the introduction of new instructional changes and according to Fish (1988), the older and less experienced teachers were, the more negatively there were likely to react. This is one of the reasons why high school teachers with a good command of English got involved in the research. They would not be hesitant about involving writing and speaking, if necessary.

4.2.2. Students

The first group of 30 students are those who will be attending high schools as freshmen in 2019 and they are the first high school students who experienced the first FSS, which must have experienced more ups

and downs of the system itself relatively to those following years, back in 2016. The 30 participants will be selected in a random manner and any attempts will not be made to select participants with either good scores or low scores, which might affect the reliability of the research. It is clear that they have one less semester during which they were pushed to study in a traditional manner less than their predecessors were.

They will take an English test published by Seoul Metropolitan Office of Education conducted in March, 2018, which were taken a year ago by students who had not experienced the FSS, i.e. *not testing* period. The test consists of listening, grammar, vocabulary, reading. It will be taken by the participants on March, 2019, with a view to giving them the same amount of time as their predecessors to improve English ability during winter season. It will be made sure that they will be controlled not to be exposed to the question items contained in the test.

The second group of students include all the freshmen in high schools in Korea back in 2018, who took the English test. They had full 6 semesters when they were exposed to traditional education system where they were forced to learn grammar, memorize vocabulary, how to read passages as fast and precise as possible. They would not have to participate in the research even though would have been possible for them to due to the fact that their results were made public back in 2018. The results refer to the outcome of six full semesters of English education.

The results of the two groups will be compared with section by section each other. The difference will be interpreted by means of the multi-facet Rasch Model. The result that are to be obtained by the comparison is highly likely to have meaningful messages to the society, regardless of its significance.

5. Conclusion

The research only contains the anticipated conclusion, which is its limitation. Still, it holds importance to a great extent.

5.1. When it comes to the results of the student group

If the difference of the results from the two groups, one under the influence of FSS and the other not, reveals no significance, it means whether *testing* exists or not does not make much difference to language education in Korea. To be more specific, students is believed to remain motivated even in the absence of *testing*. Then, what would be the reason of students continuing to hope for studying English as if they had examinations and there were a factor mediating *washback* being produced. Unlike dominant opinions referring to teachers as the main factor for *washback*, would there be any other more dominant factor?

The research intends to focus more on factors that keep students motivated to study with *not testing*. It is not difficult to come up with the reasons: they want to study in universities considered to be more prestigious than others and they want to be part of more promising workplace. Therefore, the research suggests that as long as the social value that is shared in a certain country or a context remain unchanged, that is, they want to possess as good a school degree as possible and they want to be as financially affluent as possible by getting a better job, then no matter what happens, the students will remain motivated to push themselves towards the goals. It may indicate that *prestige* and *macro-context factors* are more prominent ones than *personal factors* including teachers.

If the difference is found to be significant, it is expected that the scores from students without FSS are likely to be higher than those from students with FSS, indicating that *not testing* situations have failed to

keep students motivated in the same level at which they would have studied with *testing* playing an important role in their curriculum. In the case of Korean educational innovations like FSS, FYS, and the *revolutionary schools*, many parents have shown concerns that the schools would not finish the material that students should learn to get good scores for the university entrance exam. Moreover, it is shown that the concept is susceptible to manipulative interpretation by students, which leads to difficulty controlling students. This finding may stress the importance of *testing*, and the limitation of replacing it with assessment and evaluation of students.

Whatever result may be obtained, it is clear that *testing* itself plays such an important role in the society as a means of measure the educational progress of students, and that as long as *macro-context* or *prestige factors* like the society do not change, students will stay motivated to devote their time and energy to studying for the future.

5.2. When it comes to the results of the teacher group

The research will use the questionnaires and interview to get the answers to two questions. The first is to find out what the teachers in middle schools feel and think about the situations without *testing* and with any efforts to replace it with other methods like assessments or persuading students to partake in activities that policy-makers think helpful to improve talents, creativity, and critical thinking. It is expected that the pros and cons may come with very different opinions. They are likely to be related to students' behaviors, motivation, passion, etc., which are so intangible that it is hard to judge the current method is better than the old. The answer may be more materialized as of 2019, when the first students with FSS proceed to high schools.

The second is to see if teacher would be willing to adapt to the introduction of new instructional practices if it were not for *testing*, and to figure out that if they are reluctant it would be due to their lack of

language proficiency that ground them to halt. There is some research that argue that teachers are the prominent factor mediating *washback* being generated. And yet, this case might be a bit different. If *testing* gave way to other means like assessment and evaluation in order to give students more opportunities to do things that would not otherwise be done with *testing*, teachers may be under pressure of having to adapt to modification. Therefore, the situation of getting rid of *testing* is expected to give much clearer answer about teacher factor about *washback*.

6. Discussion

The research contains some limitations, the biggest of which is the fact that it failed to conduct the experiment to answer the research questions. The experiment is scheduled to be conducted on March, 2019, and until then it is impossible to set up the situation without getting rid of *testing* that is closely related to future high-stakes tests. The further research will help raise the reliability of the research. Also, if possible, the argument may have stronger voice if the student group can gather more than 30 participants to increase the representativeness of the group.

Among many negative *washback* of testing, it is recommended that further research attempts to see if test-takers would be less likely to depend on memorization but for *testing*? The notion is underlain that if *testing* vanishes, students will be freed from the pressure of *testing*, which leads them to rely more on critical thinking and brings about more creativity. It is worthwhile to find whether the notion is correct. It would also be worthy of conducting the research about other high-stakes tests like TOEFL, TOEIC, and IELTS. Having said earlier that it is not difficulty to find alternatives for each, each of the tests has specialties and covers expertise that may sometimes overlap with that of one another but other times do not. Last but not least, *not testing* may

arise the polarization of the status of results, which means that if you remain motivated even in the absence of *testing*, your scores will be similar with those that you would get with *testing*. However, there is no strong controlling device like *testing* for students who are not motivated for whatever the reasons might be. This might be revealed as serious issues in the near future. Hence, future research might find it worthwhile to investigate the polarization that might surface in the near future with *not testing*.

References

- Alderson, J. C. (1986). Innovations in language testing. In M. Portal (Ed), *Innovations in language testing: Proceedings of the IUS/NFER conference* (pp.93-105). Windsor: NFER-Nelson.
- Alderson, J. C., & Wall, D. (1993). Does washback exist? *Applied Linguistics*, 14, 115-129
- Bray, M., & Steward, L., (Eds.). (1988). *Examination systems in small states: Comparative perspectives on policies, models and operations*. London: Commonwealth Secretariat.
- Chapman, D. W., & Snyder, C. W. (2000). Can high-stakes national testing improve instruction: Reexamining conventional wisdom. *International Journal of Educational Development*, 20, 457-474.
- Cheng, L. (1999). Changing assessment: Washback on teacher perspectives and actions. *Teaching and Teacher Education*, 15, 253-271.
- Ebel, R. L. (1966). The social consequences of educational testing. In A. Anastasi (Ed.), *Testing problems in perspective* (pp. 18-29). Washington, DC: American Council on Education.
- Elton, L., & Laurillard, D. (1979). Trends in student learning. *Studies in Higher Education*, 4, 87-102.
- Fish, J. (1988). *Responses to mandated standardized testing*. Unpublished doctoral dissertation, University of California, Los Angeles.
- Madaus, G. F. (1985). Test scores as a administrative mechanisms in educational policy. *Phi Delta Kappa*, 66, 611-17.

- Madaus, G. F. (1988). The influence of testing on the curriculum. In L. N. Tanner (Ed.), *Critical issues in curriculum: Eighty-seventh yearbook of the National Society for the Study of Education* (pp.83-121). Chicago: University of Chicago Press.
- McEwen, N. (1995). Educational accountability in Alberta. *Canadian Journal of Education*, 20, 27-44.
- Messick, S. (1996). Validity and washback in language testing. *Language Testing*, 13, 241-256.
- Noble, A. J. & Smith, M. L. (1994). *Old and new beliefs about measurement-driven reform: 'The more things change, the more they stay the same'* (tech. Rep. 373). Tempe, AZ: Arizona State University, Center for the Study of Evaluation.
- Smith, M. L. (1991). Put to the test: The effects of external testing on teachers. *Educational researcher*, 20(5), 8-11.
- Wall, D. (1997). Impact and washback in language testing. In C. Clapham & D. Corson (Eds.), *Encyclopedia of language and education: vol.7. Language testing and assessment* (pp 291-302)
- Watanabe, Y. et. al. (2004). *Washback in Language Testing: Research Contexts and Methods*. Routledge.
- Wiseman, S. (Ed.) (1961). *Examinations and English education*. Manchester, English: Manchester University Press.