
Washback of the performance-based test of reading for EFL learners

Wiyaka

Department of English Language Teaching, Universitas PGRI Semarang, Indonesia

Correspondence

Email: wiyaka@upgris.ac.id

Abstract

The effect of tests on the teaching and learning, generally known as washback, has been recognized by scholars for a long time. However, studies on washback are usually addressed for high-stakes testing. This study investigates the washback effect of low-stakes test namely performance-based test used in measuring reading skills. Such a test is low stakes because it does not bring about serious consequence on the part of the students. The subjects of the research were 10 English teachers and 50 students of junior high schools in Semarang, Central Java, Indonesia. Data were collected through semi-structured interview and a questionnaire for both the teachers and students. The results showed that performance-based test gives positive effect in reading for the EFL learners in the areas of: students' enthusiasm in reading, reducing boredom in reading, students' curiosity on reading text content, and students' improvement on higher-order thinking skills. To the teachers it affects the areas of teaching methods, teaching materials, and time allotment.

Keywords: *Washback effect, performance-based test, reading.*

Received:

16 January 2020

Revised:

25 July 2020

Accepted:

7 August 2020

Published:

31 August 2020

INTRODUCTION

One of the changes in Indonesia latest curriculum deals with the assessments of the learning outcomes in that there is a shift from a product-oriented test toward process-and-product based tests using authentic assessments. Mueller (2005) explains that authentic assessment is a form of assessment in which students are asked to perform real-world tasks that demonstrates meaningful application of essential knowledge and skills. The importance of such assessment of real-life-situation-based performance becomes more crucial in identifying students' language abilities (Ataç, 2012). English teachers can use some methods of authentic performance-based assessment such as story retelling, oral presentation, or dramatic reading to assess students' learning outcomes.

Awareness of the use of authentic assessment is resulted from the fact that the traditional assessments so far could not demonstrate complete indicators of the real learning outcomes of the school leavers. Traditional assessment, especially in the format of multiple-choice test, is still used due to

practical consideration. For example, a study on English national examination conducted by Sabrina (2016) revealed that the questions in the examination were proven to be practical as all of them were in multiple choice format. Sumardi (2017) asserts that multiple choice test is considered practical as a more accessible assessment technique for English language teachers because it is easy to correct and to administer in a large class.

However, multiple choice techniques which are commonly employed either in the classroom tests or in high-stakes testing such as the National Examination contain some shortcomings, and among others is its incapability to capture the full range of higher-order thinking skills which are considered important in the present curriculum. As stated by Abedi (2010), traditional assessments do not afford an opportunity for students to present a comprehensive picture of what they know and are able to do in content areas. Multiple choice tests are generally incomplete since they portray an individual at a single moment in time within particular context (McTighe & Ferrara, 2011). In relation to this, Shirzadi and Ameria (2020) remind teachers not to use a multiple-choice test when the learning outcomes are addressed to communicative skills. The result of their study in two private language institutes in Mazandaran, Iran proved that multiple choice test did not give any effect on the students' productive skills.

Meanwhile, in a point of fact, National Examination belongs to high-stakes test which means it can determine the students' future (Indrawati, 2018). So, it is unfair to justify the students' success or failure only by looking at the result of multiple-choice test. Similarly, Fatonah et.al (2013) even claim that assessing the performance through written test is invalid because it does not measure what should be measured.

Research has shown that there is a very little relationship between the test scores to measure students' ability to use knowledge and skills in practical the context of laboratory work or performative work. That is why the increasing need for methods of assessments other than the multiple choice model is realistic (Bachman, 1996). There is a need to have more holistic approaches to evaluating students' learning performance as demanded by the curriculum. In the Indonesian school context with reference to the 2013 curriculum, the assessment must cover the students' language skills in using the language. In other words, the assessments should be performance-based, which means authentic assessments. Authentic assessment emphasizes the importance of the teacher's professional judgment and commitment to enhance student learning (Ataç, 2012), while Mueller (2005) asserts that authentic performance-based assessments are engaging and worthy problems or questions of importance, in which students must use knowledge to fashion performances effectively and creatively.

However, a question is raised whether such an assessment affects the teachers and students in their teaching and learning as it belongs to low-stakes testing. So far, washback is associated primarily with 'high-stakes' tests, that is, tests used for making important decisions that affect different sectors., for example, determining who receives admission into further education or employment opportunities (Chapman and Snyder, 2000).

Researches on washback usually refer to those of high-stakes tests (e.g. Cheng, 2014). It is commonly agreed that the higher the test consequence is, the more serious it will influence the students as well as the teachers in terms of their attitude, motivation, and teaching-learning process. The research on washback of alternative test is also seen as an effort to investigate whether a new or revised form of test brings about changes in teaching and learning. The change of the test type might change how the teachers perceive their role in assessment. A study by Lee (2019) reported that the change in the assessment system affected the status of English teachers in middle schools in Korea from test adopters to test developers. Such a finding is certainly promising, especially in the Indonesian context, because right now the government campaigns the implementation of authentic assessment to complement the traditional ones. Based on this assumption, then, it is important to investigate the washback effect of the government's newly-recommended type of authentic assessment.

In general, washback refers as the effect of testing on teaching and learning (Cheng, 2005; Fulcher & Davidson, 2007). In this paper, the word "test" is used interchangeably with the word "assessment" where they bear similar meaning. The role of washback is required to estimate and ensure that the test gave appropriate influence to teaching and learning in the classroom decision or in the educational system. The influence of the test is asserted as a part of classroom practice when it could direct what will happen and occur in the classroom. It includes the teaching aspects, such as the teaching technique, teaching contents, teaching material, teaching strategies, activity or time arrangement and the ways of assessing (Sukyadi & Mardiani, 2011).

This research is trying to investigate washback effects of authentic performance-based assessment particularly in reading for EFL learners in Junior High School and how authentic performance-based assessment affects teaching and learning process in reading class. In this case, retelling is one of authentic performance-based assessments implemented by the English teachers. Retelling means responding to texts based on post-reading or post-listening when the reader recalls a story in order to understand it more fully (Wiyaka et.al, 2016).

METHOD

The main purpose of the study was to investigate the washback impact of a test applied in junior high schools in Semarang City, Indonesia. This research belongs to both quantitative and qualitative methods. The study is focused on the students' attitude toward the test and the teachers' pedagogical dimensions which include teaching method, teaching material, and time allotted to reading. For that purpose, the data were collected through a questionnaire of 4-point Likert scale of agreement where: 1= Strongly Agree (SA), 2= Agree (A), 3= Disagree (D), 4= Strongly Disagree (SD). While making the analyses, the first two choices were considered as one concept indicating agreement (SA+A= A) as positive responses while the last two choices were also dealt with as one concept indicating disagreement (D+SD= D) as negative responses. Semi structured interview was also conducted to gain descriptive data dealing with the students' attitude toward the test, and the teachers'

pedagogical responses to the test. The participants of the study were ten English teachers who were selected from different school, and fifty students of junior high school from two different schools. The results of the questionnaire constituted the quantitative data which were displayed in percentages and the results of the interview constituted qualitative data which were analyzed descriptively to entail an interpretative enterprise.

RESULTS AND DISCUSSION

Washback effect of authentic performance-based test in reading

Prior to the study, the writers called the teachers in a session of coaching the implementation of retelling technique to assess students reading comprehension. This was to ensure that all research subjects used the same technique of test.

The findings showed that authentic performance-based assessments gave positive washback effect in reading for EFL learners in the matters such as: students' enthusiasm in reading, reducing boredom in reading, students' curiosity on reading text content, and students' improvement on higher order thinking skills when assessed using authentic performance-based assessments (retelling). As it was found, the English teachers used retelling as the method of reading assessment in the class as authentic performance-based assessment. The result showed that it eliminated the students' boredom in reading activity since they did not only read some texts and answer the following questions. The graphic 1 below shows a fact that retelling as method in reading assessment encouraged the students to be active readers.

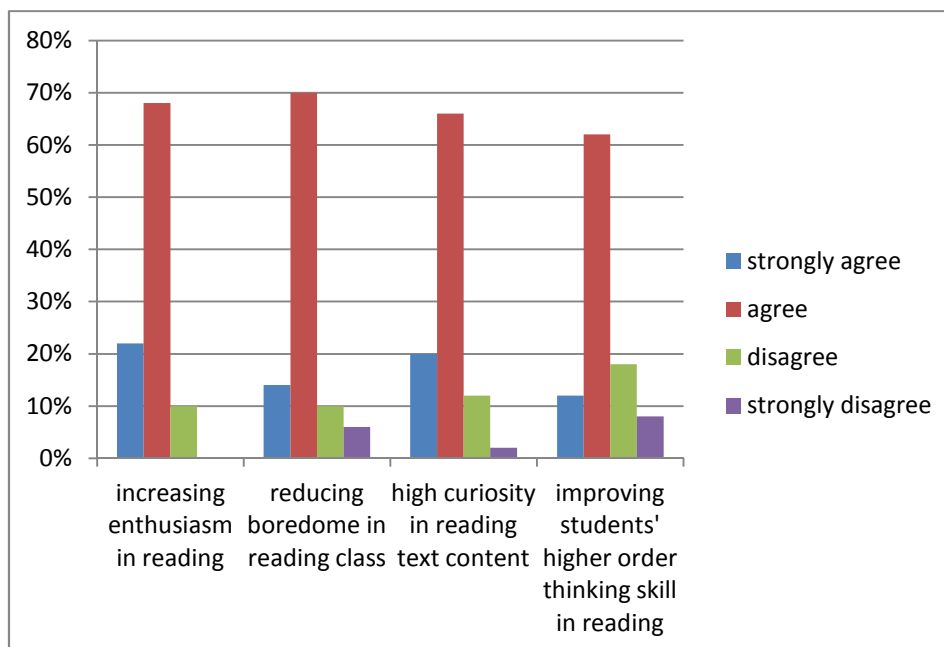


Figure 1.
Washback Effect of Authentic Performance-Based Assessment on Students' Attitudes to Reading Activity

As in figure 1 above, it shows that most students (68%) agreed that they were enthusiastic in reading class when it was assessed using authentic performance-based assessment (retelling technique). Then, 22% from 50 students strongly agreed that they have higher enthusiasm when assessed by using performance-based assessment (retelling technique). Meanwhile, 10% mentioned they did not agree on that way.

Another positive washback effect was about students' boredom which was reduced. 70% of the students agreed and 22% of students strongly agreed that they did not feel bored when assessed using authentic performance-based assessment. There were only 10% who disagreed and 6% strongly disagreed that they did not feel boring. Moreover, students also mentioned that they had high curiosity on reading text they would like to read when they were assessed using authentic performance-based assessment. 66% of the students agreed and 20% of the students strongly agreed on that statement. But 10% of the students did not agree and 6% mentioned that they strongly disagreed.

The last washback effect of authentic performance-based assessment on students' behavior in reading was about their improvement in higher order thinking skills. 62% of students agreed that they dared to retell the reading text content and gave comment relating to reading text given. Then 12% supported by mentioning that they strongly agreed on that way. On the other hand, there were 18% and 8% of the students who disagreed and strongly disagreed that they had improved their higher order thinking skills.

The findings collected from the interview reveals the same phenomena. When asked whether retelling test forced him to be active in reading the text, a student claimed that he had to read the text more intensively than before in order that he could retell the content thoroughly:

... In order to retell the text, I have to prepare it seriously. I read the story many times until I really understand the details. If I find difficult words, I ask a friend or I look up the words in a dictionary. Really, I don't want to miss any parts of the text. I list some points to help me retell the text in the correct order (Student HA).

Such a response reflects that the reading lesson his English teacher conducted so far did not generate the student's eagerness to cope with the text. The past reading activities might be monotonous and did not challenge the student to work more than just finding answers to the questions attached right after the reading materials. Dealing with the challenging activities during the performance test, another student remarked:

... The test is very challenging because in this this I and also my friends cannot do cheating anymore. There is nothing to cheat from. I think it is more about individual ability to reconstruct the story of the text (Student TA).

The other students also admitted that this test (retelling) needs more personal preparation, not only in the story mastery but also in delivering the speech. They admitted that they still had low grammar mastery, so they had to think hard how to speak in good sentences. For that reason, they had to put

serious attention to the grammar and pronunciation when carrying out the retelling test.

In relation to the reading activities prior to the retelling test, the students claimed that the reading is really fun where they learned the language features, discussed the content or the story, and shared ideas with others without necessarily fostered to focus on the reading questions. An interesting comment came from another participant:

...I like the opportunity given by my teacher. I am not only to answer the questions after reading. I have to explore the text by myself. I am free to open my dictionary, I can ask my friends, too. I imagine I can be a good story teller after I read (Student SH).

These all mean that a new type of assessment gave positive washback from the students' perspective. The use of new type of reading test has made the students change their attitude toward reading class in that reading can serve fun and challenging activities instead of boring routines.

Areas in teaching and learning process affected by authentic performance-based test

To teachers, there are three areas in teaching and learning which are affected by washback effects of authentic performance-based assessments. The areas include teaching methods, teaching materials, and time allotment. The following is the data presentation derived from the questionnaire results. The data are presented in tables describing each area of washback effects to teachers. The first area affected is in teachers' teaching method. Table 1 below shows washback effect of authentic performance-based assessment on teaching method.

Table 1.
Washback Effect on Teaching Method

No	Category	SA	A	D	SD
1	Implement activities that may promote the students' test taking skills	4	6		
2	Arrange classroom activities as usually	2	5	2	1
3	Teach test-taking strategies in the classroom meeting	3	7		
4	Select some teaching methods in a way that tend to help the students succeed on authentic performance-based assessment	4	6		
5	Select some teaching methods that tend to help develop students' skills which are more likely to be tested on authentic performance-based assessment	2	8		
6	Neglect some teaching methods that are unable to prepare the students for authentic performance-based assessment	3	5	2	
7	Teaching has been influenced by the authentic performance-based assessment	3	7		
8	Select some teaching methods without considering the kind of the assessment			8	2
9	Adjust the sequence of teaching objectives based on authentic performance-based assessment	3	6	1	

Table 1 displays that all participants used to implement activities that promote the students' test taking skills. Then, seven participants used to arrange classroom activities, but three participants did not. Moreover, they taught students with test-taking strategies, especially when the authentic performance-based assessment event was getting closer. Next, all participants selected teaching methods that tend to help and develop students' skills to succeed on authentic performance-based assessment. Surprisingly, eight participants used to neglect some teaching methods that did not give contribution to her students for facing authentic performance-based assessment, but two participants disagreed with this idea. This means that they never neglect teaching methods which are not directly related to authentic performance-based assessment. However, almost all participants confessed if their teaching methods were affected by authentic performance-based assessment. All participants disagreed with selecting teaching methods without considering the kind of the assessment. This means that the kind of assessment determines the teaching methods. This is also in line with the fact that all participants used to adjust the sequence of teaching objectives based on authentic performance-based assessment. Based on the data, the washback effects of authentic performance-based assessment on teaching methods could be categorized as a negative washback in high intensity. Cheng (2005) states that washback intensity refers to the extent to which participants will adjust their attitude to meet the demands of a test. This implies that how the teachers teach is ultimately determined by what is tested. This is sometimes misleading for some teachers in that they rely on their teaching methods on the test demands. Teaching is really test-driven in the real sense.

The results of interview reveal that generally teachers direct their teaching to what is required in the test. This means that when the test type requires students to retell what they read, they ask the students to practice similar activities during the class. Teacher AQ admitted that he was concerned with what would be tested:

...During the reading class, I ask the students to practice retelling the text. I want my students accomplish the test without any difficulties. Because the new type of test is different from the former, I need time to make students accustomed to the way the new test will be carried out (Teacher AQ).

Most of the interviewees argued that they taught the students in accordance with what the test would require students to do. They wanted to make their students successful in accomplishing the task of the performance test. The task required by the performance test was then demonstrated during the class in order that the students get accustomed to the way how the test should be accomplished. In this case, then, the lesson is narrowed to the test demand, and in turns, it may simplify the curriculum. A point of awareness dealing with the weaknesses of the new test type was expressed by another teacher:

Practicing the retelling during the reading class the whole time may make the students bored. There must be various activities to be done during the reading class to avoid the students' boredom. If the same activity is done repeatedly, it will be boring (Teacher MA).

Moreover, it is impossible to give all the students opportunities to retell due to the time constraints. Time runs very fast. In one meeting I can only invite six or seven students to perform the retelling (Teacher AW).

I think it is time-consuming. I need more time than that as scheduled if I will let each student get the turn to perform in front of the classroom (Teacher DS).

This indicates that apart from the strengths of using the performance test in reading as a technique to activate the students' language competence, there must be cautions not to bring about the students' reluctance in learning the other language skills. The teacher should be careful to give the students various activities to avoid boring routines. This is in line with Yulia and Budiharti's research finding (2019) that various teachers' questions could encourage students to partake in the classroom and at the same time could improve students' learning.

The second concern of washback is on the area of teaching materials. This deals with how the teachers prepare and use materials either from the textbook or from additional sources. Based on the questionnaire, washback effect on the area of teaching materials can be seen in Table 2 below.

Table 2.
Washback Effect on Teaching Materials

No	Category	SA	A	D	SD
10	Use materials in addition to the textbook if they will help the students' succeed on Authentic Performance- Based Assessment	4	6		
11	Give the students worksheets to review questions expected to be on Authentic Performance- Based Assessment		8	2	
12	Give the students Authentic Based Assessment exercises	2	8		
13	Choice of supplementary materials is influenced by Authentic Performance- Based Assessment	1	9		
14	Skip over certain sections in the textbook because they are less likely to be tested on Authentic Performance- Based Assessment		7	3	
15	Discuss all the materials in the textbook	4	5	1	
16	Give materials in addition to the textbook to face Authentic Performance- Based Assessment	2	7	1	
17	Display the materials of Authentic Performance- Based Assessment throughout course content	4	4	2	

From Table 2, it can be seen that all teacher participants used extra materials in addition to the textbook to help their students in reading activities. Eight participants use worksheet to review the text to see their students' comprehension, while two participants do not. In using textbook, seven participants skip over certain sections in the textbook because they are less likely to be tested in the performance test. Meanwhile, three participants did not agree with that. Another fact shows that most participants used to discuss all the materials in textbook, but ten participants do not agree. Those who agreed also mentioned that they use materials in addition to the textbook to face performance assessment. A participant even did not care about the materials prior to the test.

When interviewees were asked about the teaching materials they responded that they use materials from the textbook as well as those from other resources. Teacher AS for example claimed that he prefers using his materials of his own selection from the authentic sources:

I prefer using my own selection of reading materials. The text book does not provide various materials for reading. Performance tasks need various materials (Teacher AS).

I use the materials from the book. But it is not enough. I need different text for different student. It is ideal. But I usually divide the students into several groups. So, one group gets one topic (Teacher AR).

If the teacher is not creative in finding the authentic materials, the class will be boring and stuck because I could not improve my class (Teacher MS).

Material changing seems to be the logical consequence of using a new test type. This was admitted by Cheng (1997), in the preliminary results of a study of the washback effect of the Hong Kong Certificate of Education Examination in English in Hong Kong secondary schools, which reports that washback effect "works quickly and efficiently in bringing about changes in teaching materials.

When asked whether they give exercises similar to the test tasks they argued that they prefer to train the students with the task required by the test. Again, a teacher tends to teach what is going to be tested. Even though the test is not high-stakes, it seems that the teacher wants the students to successfully work for the test, disregarding it high or low stakes.

Textbook no longer becomes a primary source of material by teachers. When asked whether or not they still used textbook, some assert that they prefer using other materials of their own selection. If they rely on using the textbook, they could not vary the activity since all students in one class have the same text and it is not challenging. Teacher AH admitted that he selected and downloaded materials from the internet and by so doing he can freely assign different group of students with different materials.

From the viewpoints presented above, one can elicit that the teachers were forced to use extra materials in facing the performance test. Of course, this becomes a positive impact in the context of authentic assessment where

the students have to perform a certain task using different kinds of language source. This provides a clear argument that in authentic assessment such as retelling test forces teachers to be creative in preparing the materials which are supposed to be taken from authentic sources.

Dealing with the time allotment or time management, the following table summarizes the results from the questionnaire number 18, 19, and 20. Table 3 contains the detail of participant's statements.

Table 3.
Washback Effect on Time Allotment

No	Category	SA	A	D	SD
18	Do not have enough time to prepare authentic performance-based assessment.	4	4	2	
19	Spend more time to prepare authentic performance-based assessment.		7	3	
20	Do not have enough time to carry out authentic performance-based assessment.		7	3	

Based on Table 3, eight out of ten participants asserted that they did not have enough time to prepare authentic performance-based assessment, but two participants did not agree with that. To add, seven participants agreed that they would spend more time to prepare authentic performance-based assessment; while three participants said they did not spend more time to prepare authentic performance-based assessment. Next, seven of the participants felt that they did not have enough time to carry out authentic performance-based assessment, but three of them disagreed with it, which implied that they had enough time to apply authentic performance-based assessment.

CONCLUSION

From the results of findings both for teachers and students, it draws a general conclusion. Authentic performance-based assessments gave positive washback effect in reading for EFL learners in the matters such as: students' enthusiasm in reading, reducing boredom in reading, students' curiosity on reading text content, and students' improvement on higher order thinking skills when assessed using authentic performance-based assessments (retelling). The shift of assessment method influences the teaching and learning process in terms of teaching methods, material selection, and time management. From the students' perspective, some students should make themselves prepared before the test. However, they did not worry much to the test. This indicates that the performance-based test is not a serious burden to students.

REFERENCES

- Abedi, J. (2010). Performance assessments for English language learners. *California: Stanford Center*. Available at <https://edpolicy.stanford.edu/sites/default/files/events/materials/2011-06-linked-learning-performance-based-assessment.pdf> (accessed 24 October 2015)
- Ataç, B. A. (2012). Foreign language teachers' attitude toward authentic assessment in language. *Journal of Language and Linguistic Studies*, 8(2), 7-19.
- Bachman, L. F. & Palmer, A. (1996). *Language Testing in Practice*. Oxford: Oxford University Press.
- Cheng, L. (2005). *Changing language teaching through language testing: A washback study*. Cambridge: Cambridge University Press.
- Chapman, D. & Snyder, C. (2000). Can high stakes national testing improve instruction: Reexamining conventional wisdom. *International Journal of Educational Development*, 20, pp. 457-474.
- Fatonah, S., Suyata, P., & Prasetyo, Z. K. (2013). Developing an authentic assessment model in elementary school science teaching. *Journal of Education and Practice*, 4(13), 50-61.
- Fulcher, G., and Davidson, F. (2007). *Language Testing and Assessment*. Routledge: New York.
- Indrawati, T. (2018). Wash-Back of English National Examination on Teaching Learning Process. *RETAIN* 6(3), 60-68.
- Kunnan, A.J., (2014). *The Companion to Language Assessment*, First Edition. Edited by Antony John Kunnan.
- Lee, H. (2019). What are the Effects of the Change in the Assessment Systems on the Alignment between Curriculum, Teaching and Assessment? A Case in Korean Middle Schools. *The Journal of Asia TEFL* Vol. 16, No. 1, Spring 2019, 91-106.
- Mueller, J. (2005). *The Authentic Assessment Toolbox: Enhancing Student Learning through Online Faculty Development*, Vol 1, No 1. Available at http://www.google.com/vol1_no1_mueller_001.pdf (accessed 5 Dec 2015)
- Sabrina, A. (2016). A language assessment analysis of English national examination in Indonesia. *EduLite: Journal of English Education, Literature and Culture*, 1 (2), 177-196.
- Shirzadi, D., & Amerian, M. (2020). Washback effects of multiple-choice, cloze and metalinguistic tests on EFL students writing. *Indonesian Journal of Applied Linguistics*, 9 (3), 536-544
- Sumardi. (2017). Performance-Based Assessment as a Current Trend in ELT: Investigating its Washback Effects on Secondary-School Students Learning. *Kajian Linguistic dan Sastra* 2(1), 1-11. <http://journals.ums.ac.id/index.php/KLS/article/view/5347>

Sukyadi, D., Mardiani, R. (2011). The Washback Effect of the English National Examination (ENE) on English Teachers' Classroom Teaching and Students' Learning, *K@ta* Vol. 13 No. 1 98-113.

Wiyaka, Prastikawati, E.F., Sodik, J. (2016). Alternative Assessments for Reading in Senior High Schools: A Guide for Classroom Practices. *CELT Journal*, 16 (1), 49-66.

Yulia, Y., & Budiharti, F.R. (2019). HOTS in teacher classroom interaction: A case study. *EduLite: Journal of English Education, Literature and Culture*, 4 (2), 132-141.