Literature Review

Understanding Language Testing Impacts in Optimizing Its Positive Washback

[*1]Refanja Rahmatillah, [2]Rizki Fajrita, [3]Endah Anisa Rahma, [4]Veni Nella Syahputri

[*1]Universitas Teuku Umar, Indonesia, refanjarahmatillah@utu.ac.id
[2]Universitas Syiah Kuala, Indonesia, rizki.hasballah@gmail.com
[3]Universitas Teuku Umar, Indonesia, endahanisarahma@utu.ac.id
[4]Universitas Teuku Umar, Indonesia, veninellasyahputri@utu.ac.id

How to cite this article: Rahmatillah, R., Fajrita, R., Rahma, E.N., & Syahputri, V.N. (2021). Understanding language testing impacts in optimizing its positive washback. *IJELR: International Journal of Education, Language, and Religion, 3*(1), 37-42.

Abstract
This article aims to discuss the impacts of language testing on someone's personal life continuance and how language teachers should compromise to this reality and bring out the positive impacts that language testing can have on language learning (positive washback). The rationale of position of language testing in determining the flow of education, job and movement is firstly reviewed and then the cases of language testing impacts on various contexts is explored from previous studies. We then try to point out the main keys to maintain and optimize the positive washback from the existence of language testing. The consideration includes a good coordination between test producers, teachers and test-takers. Teachers, in addition, should provide follow-up teaching activities from the results of the students' tests.

Introduction
Language testing plays important roles in social life by determining the continuance of someone's education, job, and migration. Despite these crucial impacts that language testing can have, the rationale of test design received lack of attention in the past language testing studies (Shohamy, 2001). Scholars began to realize that the rationale of test design is an important aspect to be studied since most of the tests were motivated by social and political agenda (Hughes, 2003; McNamara, 2000; Shohamy, 2001). This change of language testing research focus to understand better the impacts of tests continues with the involvement of the voices of test users (test-takers, students, teachers and administrators) (Choi, 2008; Hamid, 2016; Pan, 2014; Winke, 2011), with the results of more negative attitudes towards the tests were found. Even with these negative voices, the position of language testing cannot be easily replaced because of its useful role in measuring someone's language proficiency for some critical needs (Hughes, 2003).

Considering the problem elaborated above, this article aims to enlighten how language professionals, especially teachers, can reach positive impacts from the existence of language testing for teaching and learning purposes. This objective will be accomplished by firstly describing the concept of language testing, especially about its functions and impacts. Then, some empirical research will be reviewed to understand the impacts of different language testing. Finally, some key points from the understanding of these impacts will be highlighted to show what can be done to achieve positive washback from language testing.

Knowledge of Language Testing

Language testing has been involved in decision making for some social life aspects for a long time. The position of language testing is described by McNamara (2000) similar with other testing types, such as driving and drugs testing, where they act as determinant in car driving permission or army recruitment. Language testing also acts as determinant in selecting qualified persons for education, employment, and migration purposes. Apart from this role, Shohamy (2006) argues that language testing is also utilized to enforce certain language policies. Testing is deemed as an effective and efficient tool to implement language behavior changes, such as about the priority of language skills that need to be learnt. One of the real examples for this point of view is the implementation of some language tests, e.g., TOEFL, as a university exit test. This test is applied in expectation that students will seriously improve their English ability, which is believed as a mandatory skill to be possessed in this global era, or otherwise they cannot be graduated from the university. From this implementation, it can be comprehended that language testing is a strong tool to impose language behavior changes.

Despite of this crucial role that language testing possesses, the designation of a language testing mostly involves only the people with authorities and abandons the views of the people that are intended by the test. It is as explained by McNamara (2000) that the development and organization of language testing is carried out exclusively by experts, while the affected parties can only wait for the test product results, commonly by feeling anxiety. This phenomenon happens due to the belief of traditional testing concept that the most important aspect of language testing design is to make sure that it accurately measures the knowledge intended to be assessed (Shohamy, 2001). The position of test takers in traditional testing concept is as a subject to review the quality of the tests. When the tests are viewed satisfactory, no further considerations are given about the impacts of these tests to the test takers. As a result, some test takers are questioning the correlation of the tests with their reality needs. This is what leads to the negative attitudes of test takers towards the test.

While the voices of the test takers and other parties affected by the implementation of language testing are usually ignored in the past, the recent research has taken into account this topic as the focus in language assessment field. This change occurred due to the awareness of scholars that the existence of a language test obviously impacts the individual life and society system. Brown and Abeywickrama (2010) conclude that the impacts cover both at macro level (larger contexts, including social and educational system) and micro level (smaller contexts, including individual effects). The earlier scholar's works motivate this argument, such as Chapelle and Brindley (2010), Hughes (2003), McNamara (2000), and Shohamy (2001). At macro level, McNamara (2000) for instance, highlights the impacts of language testing on test takers' socioeconomic condition by studying the policy of TOEFL as a university entrance exam. The existence of this exam pushes the prospective students to enroll in a TOEFL preparation class or hire a TOEFL instructor to help them passing the test accepting score, in which putting wealthy families in more advantageous position. At micro level, Chapelle and Brindley (2010) points out the impacts of language tests can have in teaching and learning process. The existence of tests will likely influence the way teachers preparing learning materials to fit the demands of the tests. This phenomenon can result in negative outcome that the students will not improve their language proficiency in general but rather language skills to answer tests. These two examples clearly describe how capable language testing is to make changes in the language learning behavior of a society.

Even with previous examples that language testing is more likely to result in undesirable impacts, the tests are indeed needed to screen someone's language proficiency, and thus actions to against the implementation of tests would be ineffective. The actions that need to be taken are to optimize the existence of language testing to benefit the improvement of students' language proficiency. This kind of action is actually applied in formative assessment, a term referring the use of assessment to analyze the progress of language learning in order to the students improve. Meanwhile the tests discussed here, such as TOEFL and IELTS, are categorized as a high-stake testing and summative assessment, where the information related to the test takers' detailed language proficiency is limited and thus difficult to be utilized to improve their language proficiency. Hughes (2003), however, stresses out that the scores received by the test takers should not be considered as isolated feature, rather they should be taken into account to design language learning materials. McNamara (2006), in addition, suggests that other related information to the tests should be measured carefully to help the language proficiency improvement of the test takers. Therefore, in this article, it is believed that the different impacts of language testing in different contexts should be discussed in order to extract the information needed to optimize positive washback of language testing.

Discussion
Language Testing Impacts in Various Contexts
Recent research on language testing shows a growing interest to study the voices of parties affected by the implementation of a language test (test takers, students, teachers, etc.), e.g., Choi (2008), Hamid, (2016), Li, Zhong, and Suen (2012), Ma and Cheng (2015), Pan (2014), and Winke, (2011). These studies examine various language tests, including international English language tests, such as TOEFL, IELTS (International English Language Testing System) and TOEIC (Test of English for International Communication), and other local English tests, such as CET (College English Test) and PELT (Practical English Level Test) in various contexts and places, such as schools and universities in Asia, Australia, and the USA. The examination is implemented by analyzing the perceptions of different level of stakeholders, such as test takers, students, teachers, principals, school administrators, and test administrators, towards on the effects of the tests.

Most of the participants in these studies perceive that the existence of language testing successfully brings a change for the nature of language teaching and learning process. The change, however, mainly applies to the content of teaching without actually reforms the methods of teaching, as pointed out by two studies, Li et al. (2012) and Ma and Cheng (2015). Li et al. (2012) reveals the change of learning focus that occurs to undergraduate students in China as the result of the implementation of CET. 150 university students who some already took the test and others are preparing for the test are surveyed in this study and report that they spend more time on the language skills that are tested in the test, for example speaking and listening skill. In addition, they also ask their teachers to devote the learning time more on these two skills. As a result, the teachers deem the existence of CET brings bad interference for their regular teaching process. While this test successfully changes the content that the students should learn, the learning method, however, remains the same. The teachers still use books as the main source and teacher-center method with drills and memorization techniques to teach the language skills for the students, which actually contradicts the aim of CET to encourage students' English communicative skills. Ma and Cheng (2015) also point out the same finding that language testing, in this case TOEFL iBT as Canadian university admission prerequisite, direct the prospective students to change their language learning focus. The 12 Chinese students involved in this study explain that they have to enroll in test preparation courses and they are exposed mostly to the test-taking strategies and test-features, without emphasis to improve their general their language proficiency.

In addition to the change that language testing brings to the content of language learning, Choi (2008) claims that language testing also affects the goals and methods of language learning. His study covers various EFL tests, TOEIC, TOEFL and PELT, that are implemented for Korean elementary education

system. He surveys 100 5th-grade elementary students and 40 instructors and learns that the tests influence the motivation of students to learn English to only be able to pass the tests without concerning their actual language proficiency. Moreover, the multiple-choice questions as the format of the tests limit the opportunities of the students to develop their communicative skills. Hence, it is apparent how language testing can direct the way of language teaching and learning process.

The contradiction between the goals of a language learning and the tests implemented can be also reflected from the case of TOEFL policy in the universities in Indonesia. While the aim of studying English in schools is to possess communicative skills, the application of TOEFL in universities does not encourage the use of English communicative skills. It rather focuses on the mastery of passive language skills since the TOEFL format used is the old PBT one which only assess the comprehension of listening, grammatical knowledge and reading competency. Then, the recommendation delivered by Choi (2008) perfectly fits the need to solve this problematic situation caused by the implementation of language testing which is a collaboration between authorities that design or issue the policy of tests to consider the targets' language necessity.

Another crucial impact that language testing creates is exposing the test takers to feel pressure and anxiety. This impact commonly follows with the decrease of motivation to learning the language and trustworthiness to the test. Interestingly, this phenomenon frequently occurs in the studies of language testing, as found in almost all studies mentioned earlier and the recent ones. It is as shown in Liubashenko and Yashenkova (2018) that the existence of English exit exam creates anxiety for 151 Ukranian students involved in their study. The students reveal their worry in making mistakes in the test and it is apparent when the researchers observe their exam performance. The researchers argue that this condition occurs due to the crucial role of the test in determining the students' graduation and thus leads to their negative attitudes towards the exam. Furthermore, Li, et al. (2012) finds out that the negative views held by the students towards English exams can result in the frustration and the drop of their motivation to learn the language.

Besides coming solely from the experience of taking the test, the feeling of anxiety and frustration can also come from the retake policy – a policy where test takers should retake the tests when they fail – that are applied in some language tests, especially those that serve as an entrance or exit exam. The feeling caused by this policy could result in the test takers' untrustworthiness to the test. Hamid (2016) reveals this impact by investigating the implementation of retake policy in IELTS administered in Australia. 430 people who took the IELTS test question the rationale of the policy that demands them to retake all sections of the test, not the one(s) that they fail only. As a result, they doubt the reliability of the test and even accuse the test administrators to gain commercial profits from the implementation of this policy.

While it is true that the implementation of language testing can result in negative impacts as discussed above, language testing is unarguably needed to measure someone's language proficiency. In fact, language testing can also lead to some positive washback, which is beneficial for the language teaching and learning process. Winke (2011) is one of the studies that reveal this positive impact of language testing. She explores the teachers' perceptions towards the implementation of the English Language Proficiency Assessment (ELPA) in the United States and receives mostly positive responses. The teachers describe that the test helps them to detect the students' learning needs that they could not notice before. As a result, the teachers can focus on preparing learning materials that suit the students' needs.

The existence of language testing is also found to encourage the learners' autonomy to study the language tested. It is as shown in Pan (2014) that Taiwanese students increase their time to learn English as the result of the implementation of TOEIC and General English Proficiency Test (GEPT) as university exit exams in Taiwan. Moreover, the students also optimize any English materials available to them to prepare for the test. This condition is not found to other Taiwanese students that are not imposed by the test policy. Thus, it is revealed that the implementation of language testing can also result in some

positive impacts, especially in language teaching and learning process. This positive washback is probably the motive of the implementation of language testing as a determinant policy. However, as shown in previous studies discussed, this positive washback cannot always be realized, that in contradiction, result in negative impacts. Therefore, the understanding of criteria and factors that can lead the positive washback of language testing is required, while at the same time considering those that can lead to negative impacts.

Bringing Out the Positive Washback of Language Testing
Based on the previous explanation, the stakeholders including test writers, test institutions, test takers, teachers, and school or university administrators. The positive washback of the testing can be obtained by a good coordination among stakeholders. This is supported by the arguments from Hughes (2003) and Flores and Turnbull (2003) who mentioned the efforts to increase the useful effect of a language test. The intensive and shared understanding interaction among the parties would create a good language testing. According to Brown (2010), providing positive feedback to the students can improve their motivation to study as well as be useful for teacher to prepare the lesson plan. Students are given feedback in form of their strength and weakness so that it can increase the positive washback of the test.

Other previous studies about involving stakeholders in test preparation are shown in this part. A study from Saif (2006) proves that incorporation of the university stakeholders, lecturers and students give the positive washback of a high-stake English test. The results of interview among the parties to their expectation about the new test showed many positive comments. Dikilitaş, Dollar, and Mede (2018) also interviewed the expectation from university teachers in Turkey who involved in language preparation test. The result showed that teachers' involvement gives positive impacts to the quality of the test and learning process as the researchers found gap between the content and the goal of the test and the learning practices. Zou and Xu (2017) add that the good quality of the test need to consider students' need and suit the syllabus. Based on the finding, it is believed that a good incorporation among the parties in language preparation test are highly increase the positive washback.

Another aspect should be considered in the language test is support from teachers to their students. It is line with Bachman's study (2015) who states that the main function of a test is to comprehend someone's language proficiency. Papakammenou (2018) also stresses on the implementing of communicative teaching strategies to support the attractiveness of language test class and reduce their pressure in taking test. He also points out that giving the positive feedback during the language preparation class can reduce their frustration when taking the test. Besides, Ma & Cheng (2015) also adds that teachers need to correlate the goal of the test and students' real life need in order to improve the students' motivation to learn. Thus, it can be summarized that the teachers' role in preparation of language test could be helpful and give positive washback for test.

Based on the above theories and finding, involving stakeholders such as university administrators, students and teachers must be the main priority in accommodating the students' real need in mastering the language tested. The teachers, especially, should provide the positive feedback both in form of supporting the students' interest to the language preparation class and motivate students' language performance. This has shown how the impact of language testing could affect the improvement of students' language proficiency.

Conclusion
This article discusses the importance of language testing impact in optimizing its positive washback. Language testing have been used for language policy so that it can bring some consequences for social and individual context. Some studies mentioned in the article show the negative change of the language teaching and learning process and demotivate learners. However, the language testing could be as a tool to find our more students' language needs and give improvement in language proficiency. To obtain the

desired effect, a good coroporation among the stakeholders involved in language testing should be maintained. The last element that should be considered is teachers' role in facilitating the test in order to enhance their students' language proficiency.

References

Bachman, L. (2015). Justifying the use of language assessments: Linking Test performance with consequences. *JLTA Journal, 18*, 3-22.

Brown, H. D., & Abeywickrama, P. (2010). *Language assessment: Principles and classroom practices.* White Plains, NY: Pearson Education.

Chapelle, C. A. & Brindley, G. (2010). Assessment. In. N. Schmitt (Ed.). *An introduction to applied linguistics (pp. 247-267).* New York: Routledge.

Choi, I. (2008). The impact of EFL testing on EFL education in Korea. *Language Testing, 25*(1), 39-62.

Dikilitaş, K., Dollar, Y. K., & Mede, E. (2018). English language teachers' perceptions of involvement in high-stakes testing in Turkey. In D. Xerri, & P. V. Briffa, (Eds.). *Teacher involvement in high-stakes language testing.* Cham: Springer International Publishing.

Hughes, A. (2003). *Testing for language teachers* (2nd ed.). Cambridge; New York: Cambridge University Press.

Hamid, M. O. (2016). Policies of global English tests: Test-takers' perspectives on the IELTS retake policy. *Discourse, 37*(3), 472-487.

Liubashenko, O. & Yashenkova, O. (2018). Extending the scope of the English exit exam: A study from a Ukrainian Classical University. In S. Hidri. (Ed.). *Revisiting the assessment of second language abilities: From theory to practice* (pp. 415-435). Cham, Switzerland: Springer.

Li, H., Zhong, Q., & Suen, H. K. (2012). Students' perceptions of the impact of the college English test. *Language Testing in Asia*, *2*(3), 77- 94. doi: 10.1186/2229-0443-2-3-77

Ma, J. & Cheng, L. (2015). Chinese students' perceptions of the value of test preparation courses for the TOEFL iBT: Merit, worth, and significance. *TESL Canada Journal*, *33*(1), 58-79. doi: 10.18806/tesl.v33i1.1227

McNamara, T. (2000). *Language testing.* Oxford: Oxford University Press.

McNamara, T. (2006). Validity and values. In Chalhoub-Deville, M., Chapelle, C., & Duff, P. (Eds.). *Inference and generalizability in Applied Linguistics: Multiple perspectives*. Philadelphia: John Benjamins Publishing Company.

Pan, Y. (2014). Learner Washback variability in standardized exit tests. *TESL-EJ, 18*(2), 1-30.

Papakammenou, I. (2018). Washback on language skills: A study of EFL multi-exam preparation classes. In S. Hidri. (Ed.). *Revisiting the assessment of second language abilities: From theory to practice* (pp. 437-454). Cham, Switzerland: Springer.

Saif, S. (2006). Aiming for positive washback: A case study of international teaching assistants. *Language Testing, 23*(1), 1-34.

Shohamy, E. (2001). *The power of tests: A critical perspective on the uses of language tests*. Oxfordshire; New York: Routledge.

Shohamy, E. (2006). *Language policy: Hidden agendas and new approaches*. London; New York: Routledge.

Solano-Flores, G. & Trumbull, E. (2003). Examining language in context: The need for new research and practice paradigms in the testing of English-Language learners. *Educational Researcher, 32*(2), 3-13.

Winke, P. (2011). Evaluating the validity of a high-stakes ESL test: Why teachers' perceptions matter. *TESOL Quarterly, 45*(4), 628-660.

Zou, S., & Xu, Q. (2017). A washback study of the test for English majors for grade eight (TEM8) in China—From the perspective of university program administrators. *Language Assessment Quarterly, 14*(2), 140-159.