

CRITICAL RESEARCH ON THE PEDAGOGICAL, INDIVIDUAL, AND SOCIAL IMPACT OF THE TOEFL PBT INTRODUCTION AS A TESTING INSTRUMENT

Imam Munandar

STAIN Gajah Putih, Takengon, Indonesia
imameducator@gmail.com

Manuscript received March 14, 2019, revised May 22, 2019, first published May 31, 2019, and available online June 19, 2019. DOI: <http://dx.doi.org/10.22373/ej.v6i2.4547>

ABSTRACT

This research is a descriptive research by nature which tries to demonstrate various impacts to diverse communities. It gathers several important findings from previous research and draw a comprehensive outlook on how the TOEFL test has influenced the decision makers, teachers, students and outer community themselves. It is found that there has been a significant impact as the TOEFL test is administered as a test instrument worldwide. These include positive and negative impacts. For instance, the test has demanded schools to include teaching TOEFL preparation as an integrative part of syllabus. Still, students are forced to learn English intensively, but they are more inclined to attend the TOEFL class rather than regular classes. The research also shows that students prefer studying American English dialect since it is found in the test content. In wider community, the test opens job opportunities for many people especially publication industries and English private courses. However, the test also contributes to the inequality of educational opportunity between wealth and poor families.

Keywords: *English language testing; TOEFL Test; English language institution*

INTRODUCTION

A language testing instrument is used to measure the language ability of a particular person (Fulcher & Davidson, 2007). By testing, we will be able to either to rank him or decide if he has met a criterion and then pass the test. However, while a test can give us useful information and a clear picture about a particular test taker, it

is widely believed that it also has various impacts as a result of its introduction. These consequences range from internal and external impacts. The former is the impact that affects the individual test takers and teaching-learning activities in the classrooms and the latter is the impact on a wider scale of community involving industries and business. The impacts of the testing that affect those individual and community can be positive and negative. The positive impact is seen as those impacts that can improve the pedagogical teaching in a language classroom and benefit the society, while the negative consequences are those that undermine the teaching and learning activities in achieving the goal of language education and create a social problem among the community.

THE SCOPE OF THE RESEARCH

This research is purposed to look closely at the various impacts resulted from the adoption of the paper based test of Test of English as Foreign Language (TOEFL PBT) as a testing instrument in various places. The article will begin with the literature review of impact, washback and consequential validity. What follows is the introduction of TOEFL PBT as the assessment instrument. The impacts under this analysis will involve those related to the language classroom teachings (washback), with more specific review at how the testing instrument affects teaching techniques and strategies performed by teachers, as well as learning activities in the language classrooms. The impacts also will be concerned with those affecting the individual test taker. The analysis will look at how the testing instrument influences their learning strategies and affects their life opportunities and career. Lastly, the impacts of testing instrument on social community outside the classrooms will be discussed. Furthermore, I will draw together the main points of discussion in the conclusion in the end of this article.

METHODOLOGY

In obtaining and analyzing the data, this research employs qualitative method with content analysis procedure. Content analysis can be understood as a research technique used to make replicable and valid inferences by interpreting the existing material in order to gain a broader perspective and conclusion. The research focuses on content generated by another previous qualitative study carried out

elsewhere. All data gathered by means of purposeful sampling. In this type of sampling, the documents are selected based on pre-selected criteria based on research question. Those pre-selected criteria are the written documents containing; 1. Washback of the TOEFL test, 2. Impact of the TOEFL test, 3. Consequential Validity of TOEFL test. Meanwhile, this research is limited to heavily focus on the impact of TOEFL test towards teaching English, students learning, individual test takers, and wider community (society, business, and industries) which happen in some part of the world.

RESEARCH FINDINGS AND DISCUSSION

Washback, Impact, and Consequential Validity

Anderson and Hamp-Lyons (1996) point out that it is widely believed that, examinations can influence attitudes, behavior and motivation of teachers, learners and parents. This statement implies that a test can affect both internal and external aspects of language institutions. The term “Backwash” has been generally referred to the effect of a test to the language classroom. It can be defined as “the way a test affects teaching material and classroom management” (Hughes, 1989, as cited in Taylor, 2005). McNamara (1998, p. 308) defines washback as “the extent to which the test influences teaching and learning”. Spratt (2005) identifies that teaching method and students’ learning are two of several aspects of language classroom that are affected by an adoption of a test. Anderson and Wall (1993, as cited in Anderson & Hamp-Lyons, 1996) in their Washback Hypothesis posit that a test will influence teaching and learning, with more specific influences to *what* and *how* teachers teach and what and how students learn. Furthermore, washback can affect classrooms in positive and negative ways. Morris (1972, as cited in Anderson and Hamp-Lyons, 1996) suggests that a test can ensure that the curriculum is put into effect. Similarly, language testing can be a tool in the language curriculum (Anderson, 1986, as cited in Anderson and Hamp-Lyons, 1996). Meanwhile, as negative washback, teachers are more inclined to coach students to passing the test, rather than teaching them the language proficiency according to the curriculum.

At the same time, according to Spratt (2005), washback is inextricably linked to the social and educational contexts in which a test is administered. This shows

that a test may affect not only schools where teaching-learning take place, but also outer community within which the test is enacted. This is what McNamara (2000) has referred to *impact*. It is understood as the wider effect of a test on the community as a whole, including the schools. This community may involve parents, industries and businesses. However, impact of a test towards classroom has never been fixed. Spratt (2005) suggests that scholars have debated the potential impact of tests for many years. Some of them have considered that tests have the potential to bring about a favorable impact on teaching and learning (e.g. Alderson, 1986; Bailey, 1996; Wall, 2000), while Alderson and Banerjee's (2001) review articles indicated that most researchers maintain that tests have negative effects.

The term *consequential validity* is introduced by Messick (1996, as cited in McNamara, 1998) to cover all areas which are affected by a test. According to McNamara (2000) consequential validity is an aspect of test validation which includes changes that may occur as a consequence of introduction of a testing instrument. He indicates that such changes may affect what is being measured by a test, and thus inference about a test score of a test taker can be questioned. The consequential validity is more concerned with the introduction high-stake tests since they will have impacts on a wider scope of community. It looks closely at every level of community affected by the test, from classroom to broader social and economic community that may be affected by the introduction of a test.

The Profile of the PBT TOEFL as a Language Testing Instrument

The TOEFL PBT Test is "a paper-based test that measures the ability to use and understand English in a classroom setting at the college or university level" (Educational Testing Service [ETS]). The TOEFL PBT test is a proficiency test which is "designed to measure people's ability in a language, regardless of any training they may have had in that language" (Hughes, 2003, p. 11). It is classified as a paper-and-pencil based test which is also known as traditional test which test the language skills separately including listening comprehension, structure and written expression, reading comprehension and test of written English (Shu-mei, 2009). The listening section is purposed to measure the ability to understand spoken English, which is limited to 30 minutes. Furthermore, structure section is limited to 25 minutes, which

is used to measure the ability to recognize language appropriate for standard written English. Reading comprehension section takes 55 minutes which is aimed at measuring the ability to understand non-technical reading material. Additionally, writing section is limited to 30 minutes, which measures the ability to write in English (ETS). The overall allocated time for the test is 140 minutes.

The test was developed in 1963 by the National Council on the Testing of English as a Foreign Language (Wall & Horak, 2006), which is later administered by ETS and the College whose base is in Michigan, USA. Even though the test is followed by the introduction of TOEFL CBT and TOEFL IBT in 2005, the TOEFL PBT is still valid and used in wide purposes. The test is offered six times a year in 87 countries around the world. The score of the test is accepted in more than 7,500 universities and colleges in over 130 countries accept TOEFL. Besides, other agencies make use this test, which involve immigration department, and medical and licensing agencies (ETS website). The type of test is proficiency test which is used to measure the English ability of test takers who want to pursue their education in English speaking world. Furthermore, the type of the assessment task under this review is *fixed-response task* which are realized in form of *four-option multiple-choice question*, except for writing section test. The writing test is scored from band 0 to 6, while the overall score of the test ranges from 310 to 677, which exclude the writing section. The test of writing ability integrated with TOEFL PBT is called Test of Writing English (TWE), which specifically measure the ability of (1) generate and organize ideas, (2) support ideas in writing with examples or evidence, and (3) use standard written English formats (ETS). As stated above, this test takes 30 minutes for test takers to write an essay from an academic topic.

The Impact of PBT TOEFL test on teaching English

Anderson and Hamp-Lyons (1996) and Banerjee (2003) are three researchers among those who have found the negative impacts of TOEFL PBT in their classroom investigation. These findings demonstrate that teachers tend to teach the test to students. On her study on TOEFL impact, Banerjee (2003) reveals that while the TOEFL test has led some schools to make TOEFL preparation course for students, this course teaches students test-taking techniques rather than language

proficiency. Most teachers are deviant from teaching objective that have been outlined in the communicative language curriculum. Equally, Anderson and Hamp-Lyons have pointed out that the TOEFL test clearly affect the content of English teaching in schools since in many parts of countries, TOEFL is offered as a regular language classes in schools. This shows that due to the perceived importance of teaching TOEFL for students, it has had its own place in the content of syllabus in language institutions. This syllabus specifically demands teachers to have a good understanding of teaching TOEFL for students to prepare them to enter universities that require TOEFL score. As a result, teachers often seek trainings in teaching TOEFL preparation before and during their TOEFL classes were commenced. This is a positive side as they enrich their knowledge about the TOEFL test. However, following my experience as a TOEFL course participant, teachers who have been trained teaching TOEFL test tend to include the TOEFL test material in every English lesson, especially during the time when the actual test is approaching. Although teachers have been given a TOEFL preparation class to teach the test, they cannot escape from teaching it in other regular class where they are supposed to teach English lesson, not the test.

Hamp-Lyons (1998, as cited in Shih, 2007) in her investigation, more specifically found that teachers in helping the students to pass the test, taught the skills promoted by the textbooks which generally consist of test-taking strategies and mastery of language structures, lexis and discourse rather than to any EFL/ESL curriculum or syllabus that promote model of language in use. Teachers were found to teach discrete point of language in the classroom and thus it does not meet with the communicative language teaching (CLT) prescribed in the language syllabus. Relating to this, Venon, (1956, as cited in Anderson & Wall, 1993) in a long time before, has warned this potentiality. He points out that tests "distorts the curriculum" and believes that teachers tend to ignore subjects and activities which do not contribute directly to passing the test, which is a form of excessive coaching for the test. This situation then can be considered as a negative effect of the test because as being influenced by TOEFL PBT, teachers do not promote communicative language approach which requires the integrative language teaching.

Furthermore, Anderson and Hamp-Lyons (1996) report that as being influenced by TOEFL, teachers in teaching English in classroom are found to talk more and students have less time available to talk. Also, students have less opportunity to do pair work or discussion. They also found the frequent appearance of teacher dominated lengthy explanations of grammatical points, with the use of much meta-language and little opportunity for students to practice. In my personal experience, this situation is a very nature of teaching TOEFL class. Teachers seem to be more willing to teach grammatical points of the test and the class has a monologue discussion. In this teaching, teacher gives students copies of TOEFL test and let them to give responses. Later, teachers give the correct answers and check if students have responded correctly. Then, teachers explain why the options are correct and incorrect, without involving students to discourse with their peers.

Furthermore, given the fact that TOEFL PBT does not include speaking task, most teachers subjected to Anderson and Hamp-Lyons' research have negative attitude towards the test. They argue that while the TOEFL test primarily consists of discrete items focusing on language below discourse level, and occurring independent of each other, it may not suitable with discourse-based communicative teaching. They say that the test is inauthentic and non-communicative, and that teaching TOEFL is boring and fragmentary, and they resented the time pressure they feel when they teach TOEFL preparation. Apart from this, teachers also reported that the methodology in teaching English is also influenced by TOEFL. They mentioned that it was the students who want the methodology and insisted on practice tests and work on TOEFL-like items. As a result, they felt like teaching English with a narrow direction to teaching the TOEFL test, rather than teaching English proficiency as they are supposed to.

TOEFL impacts on student's learning

Anderson and Hamp-Lyons (1996) also have provided a resourceful research finding on how TOEFL affects English learning. The research is conducted in the USA which involved overseas students from China, Indonesia, Korea, Vietnam, Japan, Mexico, Venezuela and Guatemala. The finding captures a positive impact of the TOEFL test for students. They found that the typical response of these students when

asked about the way they can improve English ability, especially TOEFL, is to have American friends, going to movies, using English outside class, and reading a lot. These examples reveal a positive washback of the TOEFL test for students, by which they are actively involved in learning English with high dedication and investment of time. The test has been able to generate the students' awareness of the importance of English in accessing English speaking universities that leads them to have a strong motivation to improve their TOEFL score.

Similarly, Mehrens and Kaminsky (1989, as cited in Anderson & Hamp-Lyons, 1996) mentions that as TOEFL is such a very high stake test, students have been forced to participate in TOEFL preparation courses. The test gives psychological pressure for students. In most cases, they need to invest their time to learn and practice with the TOEFL test in privately offered courses beyond their schools in the hope to increase their TOEFL score. Cheng et al (2004) report that a larger number of students enroll in short-term courses which typically run from 2 to 15 weeks long, which is specifically designed to help them reach the required score. This situation can be viewed from two different perspectives. On one hand, it is a positive impact that those courses can increase the probability of achieving a desired score of the test, and thus they are eligible to enter the universities they want to. On the other hand, it is expected that those beyond-school preparation courses put more burden on students as they have more learning activities to follow in addition to leaning activities in their schools. Consequently, this may ruin their performance at school as they become more the test-oriented. Relating to this, Alderson and Hamp Lyons (1996, as cited in Shih, 2007) in their study indicate the negative washback of the TOEFL test to students. They report that there were many more students in the TOEFL exam preparation classes than in the 'regular' classes especially when the test is approaching. This indicates that students are more inclined to attend the class where a lesson about passing the test techniques is explicitly taught, rather than the regular class where the English as a lesson in taught.

The TOEFL test also gives the incentive for a certain dialect of English to be more dominant than others in a particular setting of community. It has led American

English dialect is preferable. Choi (2008) reports that as TOEFL are predominantly enacted as English language testing in Korea, there is a tendency among schools students to adopt American accent of English. TOEFL test consists of American dialect and readings which are related to American culture. Equally, from my personal experience, students in Aceh even misleadingly assume that American English is the only standard form and therefore it is worth studying. This view is enhanced by the fact that the TOEFL preparation course are offered in many places in the province. Moreover, most universities have established relations to American and Canadian universities from which the TOEFL are introduced to Acehnese universities. These situations have a negative implication towards teaching English as Lingua Franca (ELF) which is being campaigned for in many countries. From the ELF's perspective, this situation is not suitable in promoting English language as not to be bound to a particular country or culture.

Another drawback of TOEFL PBT test is reported by Shu-mei (2009) that shows the test's influence on students' learning. She points out that as a consequence of taking TOEFL PBT test that does not test the speaking ability, Chinese students overseas with high mark of the TOEFL PBT score have limited skills in communicative ability in English. In learning English before the test, they direct the learning outcome to pass the TOEFL test so that they can enter the international universities. The students are found to be more inclined to learn discrete point of the test which follows the test, with lots of attention is given to receptive skills of the language.

TOEFL impact on individual test takers

McNamara (2000, p.4) points out that a language test plays a vital role in many people's life, "acting as gateways at important transitional moments in education, in employment, and in moving from one country to another". As the TOEFL test is considered as a high-stake test, its impacts towards the test takers are significant. The first effect of the test is regarded with test takers' future educational careers. The test in one hand gives ways for the test applicants to pursue their study in English speaking world; on the other hand, it becomes one of the inhibiting factors for those fail the test in accessing English-mediated universities. Interestingly,

sometime the TOEFL test is also used in non-English speaking countries as a requirement in the university admission. Hankookkyoungje (2005, as cited in Choi, 2008) reported that the test serves an academic purpose at domestic institutes of learning in Korea, primarily as a requirement for admission to and graduation from colleges and universities and graduate schools. Most foreign language high schools also require candidates to submit the TOEFL score report for entrance. He further mentions that all Korean universities adopt a policy where students have to obtain certain range of scores on some of the standardized EFL tests including TOEFL to fulfill graduation requirements. Prapphal (2008) mentions that in Thailand, a student must obtain 500 TOEFL score in order to be eligible for doctoral program. Meanwhile, some universities in Aceh controversially require their students to take the PBT TOEFL test and achieve a standardized score in order to be eligible to graduate from the university. This score varies for different universities, from 450 to 500 of the overall test score.

Additionally, the TOEFL test also can affect the job opportunity of the test takers. In some places, TOEFL is used inappropriately as an instrument to measure the ability of the test taker. In Japan, Sasaki (2008) reports that people need to obtain 550 of TOEFL score when they want to be official English language teachers. In Indonesia, TOEFL is used to test English ability of participants in Public Service Entrance Examination (PSEE). If the test takers achieve a particular criterion, they will pass the test and get the job, while they are neglected for the job in case that they fail to achieve a certain score. Given the fact that the unemployment in the country is relatively high, becoming a public servant is strongly desired, which is perceived the only chance to secure a safe job, and because of that, obtaining the highest score of TOEFL is extremely determining for the test takers.

TOEFL impacts on society, businesses and industries

The social impact of the TOEFL test can be positive and negative. While the former is applied when the TOEFL gives social or economic incentive to the community, the latter cause social problems. Prapphal (2008) suggests that the test plays a role in causing a social gap among Thai students. He reports that there has been a social concern in Thailand on the issue of equality in accessing bilingual

education in the country. For students, this bilingual education is taken for improving English proficiency and achieving a particular score in English foreign language testing including the TOEFL test. However, it is found that only the children from elite families can afford this kind of education and therefore brings about a gap with those from middle class families. In line with this, McNamara (2000) has mentioned that the introduction of a test is open to the abuse of wealthier families who can afford coaching for their children to achieve the score needed to enter a university.

Meanwhile, the TOEFL test has been able to give economic incentive for many people. According to ETS, nearly a million individuals take the TOEFL test to show their English-language proficiency every year and over 25 million people from all over the world have taken the TOEFL test since it was founded. With this number of test applicants, along with the cost of TOEFL PBT test in US\$ 160 per a test taker, ETS as a business in English language testing service gain huge amount of income. Furthermore, the TOEFL test has opened jobs and market opportunities for many people. In many countries, publication industries have published TOEFL preparation materials for the test takers. This publication is done continuously to upgrade their TOEFL test materials according to the latest version of the test content. Similarly, many private English courses offering TOEFL preparation have been established. With the test center in 87 countries, it also opens employment for many people in those countries to work in testing centers.

CONCLUSION

The impacts of the introduction of TOEFL PBT test can be considered significant to language classroom and outer community. These impacts range from positive and negative consequences. The test has demanded schools to include teaching TOEFL preparation as an integrative part of syllabus. Within classroom, the material and methodology that teachers use are influenced by the test. The teachings are found to be test-oriented where students are coached how to pass the test rather than learn the language proficiency. Teachers-students interaction also tends to be monologue, with less peer discussion. From the CLT's perspective, this situation is worsened by the fact that the TOEFL PBT measures the language discretely and does not account for speaking ability of language learners.

Furthermore, students are forced to learn English intensively, but they are more inclined to attend the TOEFL class rather than regular classes. They also prefer studying American English dialect since it is found in the test content. In wider community, the test opens job opportunities for many people especially publication industries and English private courses. However, the test also contributes to the inequality of educational opportunity between wealth and poor families. Apart from this, the absence of communicative concern in TOEFL PBT has led TOEFL CBT and IBT to be proposed. These kinds of test include speaking as an aspect of the test which is expected to give positive washback to the context of teaching communicative English in schools.

REFERENCES

- Alderson, J. C., & Banerjee, J. (2001). Testing languages from researcher perspectives. *Language Testing*, 14(1), 36-50.
- Alderson, J. C., & Hamp-Lyons, L. (1996). TOEFL preparation courses: A study of washback. *Language Testing*, 13(1), 280-297.
- Anderson, J. C., & Wall, D. (1993). Does washback exist? *Applied Linguistics*, 14(2), 115-129.
- Banerjee, J. (2003). The TOEFL CBT (Computer-based test). *Language Testing*, 20(1), 111-123.
- Cheng, L., Rogers, T., & Hu, H. (2004). ESL/EFL instructors' classroom assessment practices: Purposes, methods, and procedures. *Language Testing*, 21(1), 360-389.
- Choi, I. (2008). The impact of EFL testing on EFL education in Korea. *Language Testing*, 25(1): 39-62.
- ETS. (2010). *About the TOEFL® PBT Test*. (2010). Retrieved from, <http://www.ets.org/toefl/pbt/about> on 6 October, 2010,
- Fulcher, G., & Davidson, F. (2007). *Language testing and assessment: An advanced resource book*. Abingdon, UK: Routledge.
- Hughes, A. (2003). *Testing for language teachers* (2nd ed.). Cambridge, New York: Cambridge University Press.
- McNamara, T. F. (1998). Policy and social considerations in language assessment. *Annual Review of Applied Linguistics*, 18(1), 304-319.
- McNamara, T. F. (2000). *Language testing*. England: Oxford University Press.
- Prapphal, K. (2008). Issues and trends in language testing and assessment in Thailand. *Language Testing*, 25(1).

- Sasaki, M. (2008). The 150-year history of English language assessment in Japanese education. *Language Testing*, 25(1), 63-83.
- Shih, C. (2007). New washback model of students' learning. *The Canadian Modern Language Review/La Revue canadienne des langues vivantes*, 64(1), 135–162.
- Shu-Mei (2009). EFL education in China: Testing perspective. *Language Testing*, 12(2): 21-25.
- Spratt, M. (2005). Washback and the classroom: The implications for teaching and learning of studies of washback from exams. *Language Teaching Research*, 9(5), 5-29.
- Taylor, L. (2005). Washback and impact. *ELT Journal Volume*, 59(2), Oxford University Press.
- Wall, D., & Horák, T. (2006). The impact of changes in the TOEFL examination on teaching and learning in Central and Eastern Europe: Phase 1, the baseline study. *ETS Research Report Series*, 2006(1), i-199