GAMIFICATION-BASED ASSESSMENT: THE WASHBACK EFFECT OF QUIZIZZ ON STUDENTS' LEARNING IN HIGHER EDUCATION

Muhammad Dafit Pitoyo
Sebelas Maret University, Indonesia
Email: Mr.daves.sir@gmail.com

Sumardi
Sebelas Maret University, Indonesia
Email: arif_sumardi74@yahoo.co.id

Abdul Asib
Sebelas Maret University, Indonesia
Email: abdul.asib@yahoo.com

Abstract

Assessment and technology are two components that are very important in education. Assessment is the core of education, which provides teachers much information from their students and technology helps to achieve the learning goal more easily. Teachers are able to take the advantages of the integration of technology and assessment to get really detail information of students test to create a good teaching process. However, there are not many teachers, who really consider about assessment and the washback effect on students' learning. Thus, this study aimed to explore the washback effect of Quizizz assessment platform on students' learning. To investigate the issue, the researcher used questionnaire, observation and in-depth interviews. The data obtained would be analyzed qualitatively. The result of the study shows that students were motivated and they wanted to learn more deeply after doing several gamified test with Quizizz. They were interested in elements of game such as Leaderboard, Meme, Time restriction, and Test report.

Keywords: Assessment, Technology, Gamified test, Elements of game, Washback effect.

INTRODUCTION

Technology is developed continuously from year to year to help humans to complete their work more easily. It is employed in every aspect of social life including education. The education trend in the 21st century is to change traditional classroom into classroom based online to provide students learning convenience. The traditional classroom employment is gradually decreased along with the 21st century trend. Thus, there are many teachers in many schools together with universities struggling to integrate technology in their classroom in order to achieve better learning outcome.

The technology integration can be used teaching and assessment. However, in this paper, the researcher focuses on assessment because it is highly matter to meet a successful teaching and learning process. According to Bransford, Brown, & Cocking (2000), assessment is a core component for effective learning. Teachers will always need information of their students to anticipate teaching failure. Furthermore, teachers are able to assess their students through several types of test. Assessment is a tool for teachers to collect more information about students' learning progress systematically (Linn & Miller, 2005).

Teachers offer innovative assessment as they take the advantages of technology use

such as to gamify assessment. Several teachers use gamification to create engaging test as the response of boring test which is based on paper and pencil. The idea of gamification is from game which most of people like to play. The players of World of Warcraft game spent more than 30 hours a week playing (McGonigal, 2011). In short, the engagement of game is considered high and it will be such a breakthrough to create assessment into game-like. Gamification is non-game contexts in using game design element (Pagés et.al, 2014). Thus, it is expected to enhance students' learning motivation. Gamification can also be used for the purpose of improving user engagement, motivation and instruction (Kim, 2015). In addition, the game elements which can promote learning engagements are scoreboards, personalized fast-feedback, points, badges and leaderboards (Muntean, 2011).

LITERATURE REVIEW
Assessment

One of the important things in education which support the successful teaching learning process is assessment. It is admittedly to be a source of data which will be utilized to create anticipation. It means teachers will be able to reflect their teaching process including the students learning process to achieve the learning goals. Assessment appeared in educational issue in the 20th century. There were various literatures discussed about assessment and there were lots of experts interested in the case. They agreed that assessment is the part of educational system which is not separable as it becomes the main component of education to conduct an effective learning (Linn & Miller, 2005).

Teachers should understand the progress of students' learning progress from early moment until the end of teaching learning process in some academic year. Hereby, Assessment is the way to monitor students' learning progress. Assessment is used by teacher to collect more information about students' learning progress systematically (Linn & Miller, 2005). The information is able to be collected by many ways in the form of written and oral test such as extended performance of authentic task, responses (essays), traditional paper and pencil tests, and student self-report, (Linn & Miller, 2005). In additions, teachers can also observe their students during teaching learning process and they can also assess their students incidentally or intendedly. A good teacher will never ignore his students. Thus, the teachers' attitude towards assessment influence how students identify the study content, the class, and their work (Brookhart, 1997).

There are two types of common assessment in education namely formative and summative assessment. Both types of assessment are used to evaluate students' learning progress. Nevertheless, the way to obtain the progress is different. Formative assessment is ongoing process of evaluating students' competencies and skills. Thus, the assessment is considered longer depends on the academic period given. According to Brown (2004), Formative assessment is evaluation of developing students' competencies and skills process in the form of aid to enhance students' growth and progress of learning.

While summative assessment is evaluation of students' competencies and skills in the end of unit of instruction or course. Thus, the evaluation is not ongoing like formative assessment. According to Brown (2004) Summative assessment is the way to measure what students have learnt during some period given, and usually conducted at the end of unit of instruction.

There are basic principles of assessment which are theorized by Brown (2004) as follows:
- Appropriate assessments serve in strengthening and retention of information.
- Assessments can identify areas of strength areas which need treatment.

- Assessments can offer a sense of constant approach to elements within a curriculum.
- Assessments can promote student autonomy by encouraging students' self-evaluation of their progress.
- Assessments can motivate learners to have goals.
- Assessments can evaluate teaching effectiveness.

Gamification in Assessment

Gamification is the term which came popularly in 2010 and there are many experts discuss about the use of gamification in education such as the use of gamification to engage students learning motivation. According to de Byl (2013) the popularity of gamification is indicated in 2010 by Google search tool. Gamification is like the derivation from game in which the core is the same. Game and gamification is like using the term of *play to get engagement.* Thus, both are based on entertaining principles (Kim, 2015). However, gamification and game is different in this context of education.

Landers & Callan (2011) give definition on gamification as "the use of elements associated with game such as game mechanics to an educational purposes to create more learning engagement". Additionally, Gamification is able to enhance students' engagement, motivation and instruction (Kim, 2015). The elements of game are various and those all elements will support the idea of gamification.

As gamification is utilized to be assessment, there should be any design which associate game elements, game mechanic and game dynamic into one constituent. They are the core of the successful gamification based assessment. One of examples the design came from Werbach and Hunter (2012). They explained to design gamification based assessment into three steps. The first process is to select the dynamics of the test and second process is to select the appropriate mechanic of the test. It needs to take into account that both dynamics and mechanics of the test should be matched. The last is to select the components which fit the mechanics. Dynamics, mechanics and components of game utilized in the gamification based assessment are summarized in Table.

Table 1. Design of Gamified Assessment by Werbach and Hunter (2012)

Dynamics	Mechanics	Components
Constraints	Challenge	Content unlocking
Emotions	Challenge	Content unlocking
	Reward	Badges, Achievements, Avatars Content Unlocking
	Competition	Badges, Leader Board
	Competition	Teams
	Resource Acquisition	Points, Virtual Goods
	Feedback	Points, Content Unlocking, Badges Leader Board, Levels
Progression	Reward	Badges, Achievements, Content Unlocking
	Resource Acquisition	Points, Virtual Goods
	Feedback	Points, Badges, Leader Board, Levels
Relationship	Cooperation	Teams

Components

Components are the smallest parts which directly affect the design of gamification. To integrate the dynamics and mechanics selected in this research, the following components proposed by Werbach & Hunter (2012) were used: avatars, levels, content unlocking, the leader board, achievements, virtual goods, points, teams and badges.

Mechanics

A gamified environment consists of mechanics, which are used to create player engagement and involve essential processes (Werbach & Hunter, 2012). To highlight the dynamics selected for this design, the mechanics of challenge, rewards, feedback, resource acquisition, cooperation and competition, as outlined by Werbach and Hunter (*op. cit.*), were deployed.

Dynamics

A gamified environment consists of dynamics, which are not directly included in the process, yet make it possible to look at the design from a broader perspective (Werbach & Hunter, 2012). Among the dynamics proposed by Werbach and Hunter (*op. cit.*), constraints, emotions, progression and relationships were used in this study.

Washback Effect: Positive and Negative Washback

The term, washback, generally defined as the influence of testing on teaching and learning (Turner, 2001). In addition, Washback is the tests influence on an educational curriculum and related elements such objectives and learning goals, teaching materials, learning materials, teachers and learners and procedures (Meysam & Esmat, 2015). To conclude, the effect of test can be whether the effect will be positive or negative. It is generally admitted that the phenomenon of

washback will be either positive or negative (Saville, 2009).

The washback of a test could be considered positive if it brought about beneficial and encourage desired changes (Pearson, 1988, in Cheng & Watanabe, 2004). Every teacher certainly wants any betterment in his classroom activity including the students engagement and teachers' teaching in learning process. Sumardi (2017) said that positive washback also occurs when teacher reflect on their teaching method and then improve it to facilitate students in order that they can learn better; or students is well-motivated to learn seriously to fulfil their needs.

When teachers really pay more attention to prepare lesson more thoroughly and students and teachers can achieve their goals, it is called the positive washback (Fulcher and Davidson, 2007). It is useful for teachers to take some benefits of the test like match the instruction and the test. Turner (2009) stated that positive washback effects of a test can help a teacher change or align some instruction with general concepts represented in the test. Furthermore, the test, curriculum and the teaching should be in line. Messick (1996) stated if a test has positive washback, there is no difference between teaching the curriculum and teaching to the test and curriculum and test are highly matched.

Negative washback is defined that tests which has negative influence on teaching and learning (Cheng & Watanabe, 2004). It becomes problem that need solution. In addition, Alderson and Wall (1993) referred to negative washback as the undesirable effect on teaching and learning. Furthermore, the washback effect could be considered negative if learning principles and course objectives were not reflected in a test as anticipated (Pearson, 1988). The negative effect of washback can be in the form of motivation. As students get less motivated and discouraged to learn after they have examination, it is called negative washback. (Dorobat, 2007) says that

the negative washback occurs when students are discouraged or less motivated to learn after taking the test. It is also for teachers as they do not have any betterment in their teaching.

METHOD

This research aimed at exploring the washback effects of the Quizizz platform to be the media of the assessment as the use of ICT becomes the trend of education today. This study only explored the effect of Quizizz platform for students learning motivation in English TOEFL structure class. The respondents of this research were the students of the third semester who take English Course in a university in Surakarta. There were 18 students participated in this research. Furthermore, the sample selected based on some criteria which is it is called purposive sampling. According to Arikunto (2010), purposive sampling is the method of choosing sample by taking subject not based on the level or area, but based on the particular purpose. The sample of the study were the students who got good score in English structure subject.

The empirical data were gained through observation, in-depth interview and questionnaire. The observation showed the condition of the class during teaching learning process and the questionnaire showed students' opinion about the use of Quizizz for assessment. Additionally, in-depth interview showed detail opinion of the respondents about the use of Quizizz for assessment. It was used to be a verbal justifications for the respondents for what the researcher missed as conducting observation in classroom (Widodo, 2016).

The data were analyzed qualitatively as followed with interpretative enterprise.
1. Does Quizizz platform offer positive washback effect on students' learning?
2. What are game elements which make students interested in?

FINDINGS

The Washback Effect of Quizizz Employment in EFL Classroom

The researcher used questionnaire to know the kinds of washback effect on students learning. There were eight items with Likert scale (strongly disagree, disagree, neither agree and disagree, agree, strongly agree) and the items were divided into two from (i.e favorable and unfavorable item). It was administered to know the consistency of students' preference. The findings the researcher obtained as follows:

Table 2. The Frequency of The Washback Effect of Quizizz Employment on Students' learning

Linkert scale	FREQUENCY							
	Number of item							
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Σ SA (5)	4	1	2	3	3	4	5	4
Σ A (4)	10	7	9	13	9	9	9	10
Σ NAD (3)	2	9	6	2	5	4	3	4
Σ D (2)	2	1	1	0	1	1	1	0
Σ SD (1)	0	0	0	0	0	0	0	0

The table shows the frequency of students' preference to the items which is scaled from 1 until 5 (1=strongly disagree, 2= disagree, 3= neither agree nor disagree, 4= agree and 5=strongly agree). For the following action, the frequencies of each item would be presented in mean score in order that the students' preference for each item was easily read.

Table 3. The Mean Score of Each Item from Washback Effect of Quizizz Employment

Items	Mean
Students are motivated learn Toefl Structure after doing the test with Quizizz.	3.889
Students study seriously about Toefl structure due to Quizizz employment.	3.444
Students are willing to pay attention to teachers' explanation about Toefl structure.	3.667
Quizizz makes students get a good score in Toefl Structure.	4.056
Quizizz helps students to find their needs in learning Toefl Structure.	3.778
Quizizz causes the bad students' performance in learning Toefl Structure.	3.889
Quizizz contributes to increase students anxiety in learning Toefl Structure.	4
Quizizz causes students less motivated to learn more Toefl Structure.	4

The table showed that almost all items obtained high scores. Only one item was identified under 3.5 mean score which is about the seriousness of study about Toefl structure after doing the test with Quizizz. There are two items obtaining 4 namely about students' anxiety and students' learning motivation. Additionally, the highest score is 4.056 in which students could get better score due to Quizizz employment. However, to find the clearer washback effect, the all items would be counted to know the average of mean score or it is called grand mean score. Thus, it will be

the answer of understanding the washback effect of Quizizz employment on students' learning.

Table 4. The grand mean of the washback of Quizizz employment in higher education.

SCORE							
1	2	3	4	5	6	7	8
20	5	10	15	15	20	25	20
40	28	36	52	36	36	36	40
6	27	18	6	15	12	9	12
4	2	2	0	2	2	2	0
0	0	0	0	0	0	0	0
70	62	66	73	68	70	72	72
MEAN SCORE							
3.889	3.444	3.667	4.056	3.778	3.889	4	4
GRAND MEAN				3.84			

After, all mean scores of each item were counted, the grand mean which the researcher obtained from the questionnaire administered is 3.84. It is closely to 4. It means there is a good influence of the test to students learning.

The Students' Elements of Game Preference
The data were collected by using in-depth interview and the data showed that there were several game elements which made students were interested in. Most of students were interested in time limit of test (time restriction), meme, test report, and leaderboard in Quizizz. Those features or game elements contribute to the interactive assessment in which it promotes a successful assessment. AA said

"*Quizizz creates challenging test so that all students are competing to be the best. Time restriction which is applied in the test helps me to focus more in answering the questions. Thus,*

6

the time will not be wasted. It think it is good for the students."

One of features of Quizizz which is considered unique is meme. It was showed every after students answered a question. There are two kinds of meme used in the test (i.e meme for incorrect answer and for correct answer). As students answered a question correctly the meme would be in the form of appreciation word and motivational word if the answer was incorrect. BB said

"I like the way the teacher using Quizizz especially for Meme. It makes me think positive because every after completing a question, there will be a meme which contain positive words. It is kinds of motivational word as well. Furthermore, I don't fell under pressured during doing my test."

To make the test more competitive, Quizizz provides leaderboard. The students felt that the test was like competition. They have to compete with their friends in the class. Leaderboard generally increases students' motivation and faithfully represents the grade of competition. CC said,

"I love leaderboard feature in Quizizz because I am motivated to compete to be the best among my friend in the class. I can see the rank during doing my test."

The last game element which students were interested in is test report also known as feedback. The test report will be given shortly after students complete the test and it will be in the form of document which students can download. The students can save so that they can learn from it. A student, DD said

"Test report which is given to me help me to identify weakness. I could see my answer well and I try overcome my weakness because I don't want to get bad score one more time. I decided to learn from my failure."

DISCUSSION
The Washback Effect of Quizizz Employment in EFL Classroom

The questionnaire which had been presented in finding identified several indications of positive washback effect happening in the classroom. The positive washback happened because the use of game elements integrated in the test. The washback was considered positive because it helped students to achieve their goal or objective of their study. The washback of a test could be considered positive if it brought about beneficial and encourage desired changes (Pearson, 1988, in Cheng & Watanabe, 2004). Thus, they were interested in Quizizz which used the idea of gamification. The technology can helps someone to get what is needed. The employing Quizizz in higher education was considered motivating for students because it was engaging. Higher students prefer to do computer-based test than paper based test because they obtain more benefit (Clariana & Wallace, 2002).

The positive washback effect of Quizizz on students learning was in the form of motivation. Students were motivated to learn more on subject which was tested with Quizizz. The questionnaire showed that the mean score of the motivation aspect of each item were 3.889 and 4. The mean score of 3.889 was obtained from the favorable item. It means that the item was designed with positive statement. The mean score of 4 was obtained from unfavorable item also known as negative statement item. Furthermore, both score showed the consistency that students were motivated due to Quizizz employment. As the observation, students were very active in answering teacher question during teaching learning process. The students occasionally asked the teacher as they did not understand about some materials. Motivation always influence students' learning. It is in line with Sumardi (2017) who said that motivation was an important aspect which influenced the success of students' learning.

Besides contributing to students' motivation in learning, Quizizz also helped students to focus more on their teacher

explanation. The mean score of the aspect is 3.667. They assumed that the effect of gamification made them to take more attention during teaching learning process because they wanted to enrich lots knowledge as the preparation of the next examination. Additionally, the students realized what was needed in learning. They could identify their weakness and mistakes after doing the test. They tried to make solution over the problems by analyzing the test report. That gamification feature provided the students test record including the correction of students incorrect answer to some questions. Assessment contains information about students' learning progress systematically (Linn & Miller, 2005). In addition, the result of assessment give students access to identify the study content, the class, and their work (Brookhart, 1997).

As students could identify their weakness or understand their need in learning, the students would not be suffered with anxiety. The students could enjoy their learning process. The mean score of anxiety aspect from questionnaire was 4. The item was unfavorable item. Thus, the real score was actually 2 and it was converted to favorable item to be 4. Quizizz decrease students' anxiety during the test and after the test. As result, the students assumed that could achieve a good score. It was also showed in the questionnaire item of score aspect. The score was 4.056 and it is the highest mean score of the questionnaire. In short, the data was confirmed with their assumption.

The students' elements of game preference

Werbach & Hunter (2012) state that the elements of game in gamification-based assessment such avatars, levels, content unlocking, the leader board, achievements, virtual goods, points, teams and badges. However, there many more elements of game which were created by the developer not included in the theory of Werbach & Hunter (2012). As the findings obtained, there are several game elements which students were interested in namely time limit of test (time restriction), meme, test report, and leaderboard. Students are engaged with the elements due to the function of each element. The first elements was functioned to make students more focus on their work and it prevented students to cheat during the test because they competed also with time. As time is up, they cannot answer the question and go to the next question. Thus, the test result will be representative of students' performance. A student said that time limitation challenged them because they need to be really focus on his work.

The second game element which students like is Leaderboard. It is designed to give students real-time information about students' rank during the test. It is showed to students as they answer each question. The rank is counted based on the speed and the accuracy of their answer. Thus, as two students answer some questions correctly, the score or rank will not be the same because time is included for consideration. From the interview and observation, students were challenged to compete among their friends in the class. The test became the arena of competition. It was also supported with time limitation in Quizizz. The competitive environments is created.

For enhancing students' motivation and releasing test tense or stress, meme is the one which becomes the breakthrough. It is quite unique as generally known that meme is usually used in social media. In Quizizz, the meme gave students spirit after answering some question. As they answered incorrectly, the meme would motivate them with some words not to give up. It is useful for them because in general test or examination usually much pressure and stress. Accordingly, it influence students' learning activity.

The last game element is test report. The test report gave information to students about their test. Students could reflect their learning

progress by analyzing the test report. Students had assumption that test report was really useful for them to know their weakness which was needed a treatment. This kind of element includes in feedback function because it helped students to make any learning betterment. It provides some particular information about their performance in order that they can understand whether their answer is correct or not so that students can indicate some alternative strategy to overcome their weakness (Webb, 2013).

CONCLUSION

This study revealed that Web-based gamification with Quizizz for assessment gave positive washback effect on students' learning. It happed because Quizizz provides many features or game element and the students' favorite game element are Leaderboard, Meme, Time limit and Test report. Those made students motivated. They were happy during doing the test as they play a game. Accordingly, it influenced their learning due to the Quizizz employment. They struggled to learn more after doing the test and they were helpful as they could obtain test report which contained very detail information about their test. Furthermore, they would overcome their weakness and they could obtain a good score of the test. There were several implications of the study which the researcher concluded. The students are interested in technology which serve them something interesting. The schools should improve facilities including the availability of computers and internet access so that the teachers could improve their teaching. They are able create some teaching innovation as the demand of the revolution of industry 4.0.

REFERENCES

Alderson, J. C., & Wall, D. (1993). Does washback exist? *Applied linguistics*, *14*(2), 115-129.

.

Bransford, J., Brown, A., & Cocking, R. (2000). How people learn: Brain, mind, experience and school. Washington, DC: Commission on Behavioral and Social Sciences and Education, National Research Council.

Brookhart, S. M. (1997). Effects of the classroom assessment environment on mathematics and science achievement. The Journal of Educational Research, 90(6), 323-330. https://doi.org/10.1080/00220671.1997.10544590

Cheng, L., & Watanabe, Y. (Eds.). (2004). *Washback in language testing: Research contexts and methods*. Routledge

Clariana, R., & Wallace, P. (2002). Paper–based versus computer–based assessment: key factors associated with the test mode effect. *British Journal of Educational Technology*, *33*(5), 593-602. https://doi.org/10.1111/1467-8535.00294

De Byl, P. (2013). Factors at play in tertiary curriculum gamification. *International Journal of Game-Based Learning (IJGBL)*, *3*(2), 1-21.

Dorobat, D. (2007). *The methodology of evaluation and testing*. Athena: Ministerul Educatiei si Cercetarii.

Fulcher, G., & Davidson, F. (2007). *Language testing and assessment*. London, England: Routledge.

Kim, B. (2015). . Designing Gamification in the Right Way. *Library Technology Reports*, *51*(2), 29-35.

https://journals.ala.org/index.php/ltr/ article/view/5632

Landers, R. N., & Callan, R. C. (2011). Casual social games as serious games: The psychology of gamification in undergraduate education and employee training. In Serious games and edutainment applications (pp. 399-423). Springer, London. Springer, London. https://doi.org/10.1007/978-1-4471-2161-9_20

Linn. R.L. and Miller, M.D. (2005) Measurement and Assessment in Teaching (9th edition). Englewood Cliffs, NJ: Prentice Hall.

McGonigal, J. (2011). *Reality is broken: Why games make us better and how they can change the world*. Penguin.

Messick, S. (1996). Validity and washback in language testing. *Language Testing, 13*, 241–256. https://doi.org/10.1177/02655322960 1300302

Michael Miller. 2014. My Facebook for Seniors. USA: Pearson Education.

Muntean, C. I. (2011, October). Raising engagement in e-learning through gamification. In *Proc. 6th International Conference on Virtual Learning ICVL* (Vol. 1). sn. http://icvl.eu/2011/disc/icvl/documen te/pdf/met/ICVL_ModelsAndMetho dologies_paper 42.pdf

Saville, N. (2009). Developing a model for investigating the impact of language assessment within educational contexts by a public examination provider. The University of Bedfordshire. http://uobrep.openrepository.com/uo brep/handle/10547/134953

Sumardi (2017). Performance-Based Assessment as a Current Trend in

ELT: Investigating Its Washback Effects on Secondary-School Students Learning. *Kajian Linguistik dan Sastra, 2, 2-11*. https://doi.org/10.23917/kls.v2i1.534 7

Turner, C. (2001). The need for impact studies of L2 performance testing and rating: Identifying areas of potential consequences at all levels of the testing cycle. In *Experimenting with uncertainty: Essays in honour of Alan Davies* (pp. 138-149). Cambridge University Press.

Turner, C. E. (2009). Examining washback in second language education contexts: A high stakes provincial exam and the teacher factor in classroom practice in Quebec secondary schools. *International Journal of Pedagogies and Learning*, *5*(1), 103-123. https://doi.org/10.5172/ijpl.5.1.103

Webb, E. N. (2013). Gamification: when it works, when it doesn't. In *International conference of design, user experience, and usability* (pp. 608-614). Springer, Berlin, Heidelberg. https://link.springer.com/chapter/10. 1007/978-3-642-39241-2_67

Werbach, K., & Hunter, D. (2012). *For the win: How game thinking can revolutionize your business*. Wharton Digital Press.

Widodo, H.P. (2016). *Engaging young learners of English in a genre-based digital storytelling project*. Cambridge: Cambridge University Press.