



The Impact Of Undergraduate Program Assessment Methods On Bs English Students In Pakistan

Muhammad Umer, Assistant Professor, Islamia College Peshawar, Pakistan.

Email: mumer@icp.edu.pk

Dr. Khalid Azim Khan, Assistant Professor, Deanship of University Development and Quality Assurance, Umm Al-Qura University, Saudi Arabia. Email: kakhan@uqu.edu.sa

Abstract

This article reports the impact of the university's undergraduate assessment system on learning and teaching. Undergraduate programmes in Pakistan heavily rely on teacher assessment of students' learning outcomes during a semester. However, teachers' formative assessment practises appear to be in contrast with what formative assessment literature suggests. The purpose of this research was to study how university undergraduate assessment methods affected learners in a public sector university and how teachers used assessment as a teaching tool. Data were collected from the university English-major learners of four semesters—1s, 3rd, 5th and 7th. The data was subjected to both quantitative and qualitative analysis. The findings showed that the assessment system had a partially positive washback on students' learning because it discouraged rote learning of materials to some extent, helped students practise their IT skills, and developed their self-assessment skills. However, the teachers did not align curricular tasks with the assessment practices, and most teachers did not give formative feedback to students. Therefore, it is strongly suggested that teachers' assessment literacy be developed.

Keywords: university teaching, undergraduate learning, assessment methods, washback

1. Introduction

Assessment has always been an integral part of teaching and learning. It provides information which is used for a variety of teaching, learning, and administrative purposes. It is used to improve the quality of learning, achieve the learning outcomes, and motivate the learners. Moreover, it also assists teachers to revisit their teaching methodology and even the course contents if needed. Assessment is conducted in various ways, in the form of tests, quizzes, assignments, presentations, and final examinations, enabling teachers to measure their learners' achievements and further learning needs. The teacher's role is central in all the assessment methods. This increased responsibility requires teachers to shift from deciding whether students passed or failed to modifying classroom activities and

instruction for maximizing learning (Rabinowitz, 2010). It is suggested that assessment should be an ongoing process rather than an isolated event. For this reason, formative feedback becomes a crucial ingredient in assessing students' English. Instant feedback can prevent students from repeating mistakes and errors. However, it is widely acknowledged that feedback is only effective if it is timely and clearly articulated in order to enable learners to overcome their shortcomings (O'Malley & Pierce, 1996).

Language assessment influences language learning and teaching, a phenomenon known as the 'washback effect' in language testing literature. Such an effect is generally perceived to be either positive, having beneficial effects on teaching and learning, or negative, having detrimental effects (Alderson & Wall, 1993). A positive impact on learning could be achieved through quality assessment methods that are valid and reliable and accurately reflect students' learning. Similarly, it involves feedback as one of its important factors. Feedback enables learners to know about their weaknesses as well as their improvements, which serves as a booster for their motivation and confidence.

The relevance of the study is entailed due to the nature of the country's education system, which is mainly examination-driven. The examination determines what the teachers teach as well as how they teach (Alderson & Wall, 1993). Likewise, the examination also determines the students' strategies for learning. However, despite the significance, very few research studies regarding the washback effect have been conducted at university level in Pakistan. Hence, the current research study assesses the washback effect at university level. The following two questions guided this research:

- a. How do undergraduate assessment methods affect students learning in public sector university in Pakistan?
- b. Do teachers use assessment instruments for improving students' learning in undergraduate programs in Pakistan?

2. Literature Review

Washback is generally defined as the effect of testing on teaching and learning, which can be harmful or beneficial (Hughes, 2003). Earlier, the term 'backwash' was used to refer to the way a test affects teaching materials and classroom management (Briggs, 1995; Hughes, 1989). However, today, in the field of applied linguistics and the language testing community, the term 'washback' is commonly used (Weir 1990; Alderson & Wall, 1992; Alderson, 2004). Washback is also referred to as measurement-driven instruction (Popham, 1987), which means, as Messick (1996) maintains, "the introduction and use of a test influences language teachers and learners to do things they would not otherwise do that promote or inhibit language learning" (p. 217). Washback is also called curriculum

alignment (Smith, 1991; Resnick&Resnick 1992). It is assumed that the content and format of the test should overlap with the content and format of the curriculum (Cheng & Curtis, 2008) so as to steer teaching in the desirable direction. Bachman and Palmer (1996) regard washback as a feature of a wider phenomenon known as test impact. They suggest that test impacts should be viewed both in terms of their micro and macro effects, which are not only in a classroom but on educational systems and societies as well.

Generally, there are two major types of washback effects: either positive or negative, based on their beneficial or harmful impact on educational processes (Hughes, 1989). Supporting the view, Xu and Liu (2018) have argued that:

A test, if well designed and appropriately used, may facilitate teaching and learning, providing useful information and generating strong motivation. This is positive washback. On the other hand, if poorly designed and inappropriately used, a test may lead to negative washback. (p. 19)

For instance, if a learner is asked to write an essay to examine his writing ability but the question given is based on history or biology, which the learner does not have sufficient knowledge about, then this may lead to a negative washback effect (Harmer, 2007). Similarly, a test can have a negative impact in a case, as Alderson (1992) says, when anxiety is generated in the students or "concern in the teachers" about the consequences that will follow the poor performance of the students. Further, he says that this argument is based on the logic that students perform abnormally under pressure and teachers' embarrassment due to the fear of embarrassment, which is linked to the students' poor results, leading to "teaching to tests" by narrowing down the curriculum (Alderson & Wall, 1992).

The influence of the tests on teaching and learning is so remarkable that some researchers have even advised that "a test's validity should be measured by the degree to which it has had a beneficial influence on teaching" (Alderson & Wall, 1992,pp. 3-4). Therefore, it is preferable to take washback and validity in association "to denote the quality of the relationship between a test and associated teaching" (Alderson & Wall, 1992, p. 4). Morrow (1986) uses 'washback validity' to indicate washback is a part of validity and advises researchers to go to the classrooms to observe how beneficial washback is achieved when examining validity. Messick (1996) regards washback as a consequential aspect of validity and emphasises that washback should be "gauged by good or bad practise that is evidentially associated with the use of a test" (p. 245).

However, contrary to the notion of washback validity, some researchers argue that there are other factors, not just the design of the tests, that could affect the presence or

absence of a washback effect (Tzagari, 2009). In this regard, Ferman (2004) proposed that it is difficult to see a direct connection between washback and validity because validity is not a property of the test, but rather the meaning of the test scores. Due to the complexity of washback, Alderson and Wall (1993) said that it was impossible for validity and washback to be directly related. Davies (1997) suggested that "the apparent open-ended offer of consequential validity goes too far. It is not possible for a tester as a member of a profession to take account of all possible social consequences" (p. 335). Messick (1996) as well as pointing out the complexity, since washback involves "a function of numerous factors in the context or setting and in the people responding as well as in the content and form of the test," recommended "rather than seeking washback as a sign of test validity, seek validity by design as a likely basis for washback" (pp. 251-52).

Despite the contrasting opinions stated above, researchers agree on the necessity of investigating the effects and consequences of tests (Alderson & Wall, 1993; Alderson, 1995; Messick, 1996). Researchers have given different models of the mechanism of washback. However, it is unclear which one can best depict the washback phenomenon and illustrate its working mechanism (Xu & Liu, p. 33, 2018). Hughes (1993) presented a trichotomous model for the mechanism of washback. He proposed that it be built around three components: participants, process, and products. Participants are students, classroom teachers, administrators, materials developers, and publishers, whose perceptions and attitudes toward their work may be affected by a test. A process is any action taken by the participants which may contribute to the process of learning. Products are what is learned (facts, skills, etc.) and the quality of learning. Bailey (1996) also put forward a basic model to show the operation of a washback. Bailey's model is more comprehensive. It encompasses not only the participants, processes, and products, but also their interactions and reactions to the test (Ren, 2011). She differentiated "washback to learners" from "washback to program" which is "results of test-derived information provided to teachers, administrators, curriculum developers, counselors, etc." (Bailey, 1996, p. 264).

Alderson and Wall (1992) in their seminal work "Does Washback Exist?" gave a comprehensive construct of fifteen hypotheses about washback which ranges from most general to somewhat refined (p. 9). Their model includes washback hypotheses associated with teaching (1, 3, 4, 7, 9 and 11), learning (2, 5, 6, 8, 10 and 11), content (3 and 5), method (4 and 6), practise (3-10), and attitude (11). Furthermore, the last four hypotheses (12-15) were related to the differences between tests and participants (Xu & Liu, p. 25, 2018). Similarly, Green (2007), during his research of washback on IELTS, presented a model comprising two dimensions, namely direction and intensity. According to Green (2007), the direction of washback is related to the overlap between test constructs and test characteristics, while the intensity of washback effects is determined by test stakes

and test difficulty. Moreover, the washback effects of the same test may vary from participant to participant because of their different experiences and beliefs (Xu & Liu, 2018, p. 26). Empirical studies conducted so far strongly recommend further in-depth investigation into the phenomenon in contexts across systems (Watanabe, 2005; Cheng, 1999 & 2004; Green, 2007; Ferman, 2004; Aftab, Qureshi & William, 2014; Ahmad & Rao, 2012; Shamim, 2011).

1. Methodology

This research was a survey of students' opinions about the association between how they learnt and how teachers assessed their learning and how their teachers utilised assessment to improve the students' learning. Because of the COVID-19 pandemic, the researcher could not collect the data manually. Therefore, the researcher sent an online link to the respondents through email and other social media sources in order to reach out to as many respondents as possible. The research instrument used was an online questionnaire adapted from (Umer, Javid, & Farooq, 2013). The data collected through the questionnaire was quick to analyze.

The 35-item questionnaire dealt with every aspect of the washback effect on teaching and learning strategies. The completion of the questionnaire took approximately 30 minutes. The questionnaire was created using Google Forms. The Google Forms service enables researchers to collect information through a link or an email. Furthermore, it automatically saves the responses in a Google Sheet. All the information can be collected in this single place. From there, the information can be easily exported to other software such as Excel and SPSS for analysis. The questionnaire form contained closed-ended and open-ended questions to collect both quantitative and qualitative data.

Apart from collecting demographic information of the respondents, there were 33 closed-end questions and 2 open-ended questions that asked for the opinions of the respondents. The questionnaire was on a five-point Likert scale (Likert, 1932), ranging from "strongly agree" to "strongly disagree" (strongly agree; agree; neutral; disagree; strongly disagree). Closed-ended questions were used because they are advantageous in terms of time and money; they are easy to analyse through software such as Excel and SPSS in a short period of time, and the quality of the questionnaire can be judged by three major standards: validity, reliability, and practicality. However, to get a holistic and in-depth view of the attitudes and perceptions of the respondents about their assessment system, open-ended questions were also included at the end of the questionnaire.

Before the collection of main survey data, the researcher conducted a pilot study (Dawson, 2002). The questionnaire was sent to students who were not part of the population. Eight responses were received, which included five female and three male

respondents. After the collection of the pilot study data, the data was analysed for the reliability of the instrument. For this purpose, Cronbach's Alpha was calculated (Quansah, 2017) to be 0.89.

Results

Table 1.2 Description by Gender

	N	%
Female	57	56.4%
Male	44	43.6%

Table 1.3 Description by Semester

	N	%
2 nd Semester	20	19.8%
4 th Semester	23	22.8%
6 th Semester	29	28.7%
8 th Semester	29	28.7%
	101	100%

Analysis of the Questionnaire and Presentation

The five-point Likert scale is considered an interval scale. The meaning is very significant. From 1 to 1.8, it means "strongly disagree," and from 1.81 to 2.60, it means "disagree." From 2.61 to 3.40, it means neutral. From 3.41 to 4.20, it means agree, and from 4.21 to 5, it means strongly agree (Pimentel, 2010). In this research, strongly agree and agree are considered as agree, and likewise, strongly disagree and disagree are taken as disagree only, in order to make the analysis easily perceptible. The items of the questionnaire are presented in tables 1.4, 1.5, and 1.6 and then analysed as follows:

Table 1.4 Items that the Respondents Agreed with

Items	N	Mean	SD
1. I discuss the topics of my assignments with my friends and teachers to understand them well.	99	3.96	.925
2. I study several sources or books to write my assignments.	101	4.19	.902
3. I study several pages of a single course book or a website to write my assignments.	101	3.65	.994
4. I get penalty if I don't submit my assignment on or before the	101	3.45	1.100

submission deadline.			
5. I strictly observe the deadlines of assignment submission.	101	4.41	.777
7. I write assignments for all courses with clear introduction, body and conclusion in my own words.	101	3.99	.866
8. I write assignments using my own words instead of copying the content from a book or the internet.	101	3.76	.929
9. My all assignments include more than 1000 words.	101	3.56	1.014
10. The grades of assignments are important for me to pass my courses and score good GPA.	101	4.31	.703
19. I study more than one book or source to understand a topic very well.	101	4.23	.691
24. I like the kind of quizzes or exams where I use my own ideas, experience and knowledge.	101	4.36	.687
26. Assignments, quizzes and midterm exams help me know my progress on each course and how to improve myself.	101	3.96	1.009
30. My mistakes in a quizzes, midterm exams and in assignment help me avoid those mistakes in next quizzes, exams or assignments.	101	4.06	.978
32. My teachers don't mind when I try to get their feedback on my quizzes, exams or assignments and to overcome my weaknesses.	101	3.45	.995

According to the above data, most of the students agreed that they discussed the assignment topics with their friends and teachers, as well as studied more than one book or source for understanding a topic. Similarly, for writing their assignments, they studied several sources or books and several pages of a single course book or website. In addition, most of them used their own words instead of copying them and observing the academic rules of writing. Likewise, they confirmed that they strictly observed the deadlines for assignments because, in other cases, they got a penalty if they did not submit the assignment on or before the deadline. Furthermore, they wrote more than 1000 words in their assignments and mentioned sources while quoting text in the assignments. Furthermore, they enjoyed the kinds of quizzes where they used their own ideas, experience, and knowledge, and these assignments, quizzes, and mid-term exams helped them improve their mistakes and made them avoid those mistakes in future quizzes and exams.

Table 1.5 Items that the Respondents Disagreed with

Items	N	Me	SD
-------	---	----	----

		an	
11. When I copy the content of my assignments from the internet or course book I get good grades.	101	2.56	.953
14. All of my teachers tell me about the mistakes in my assignments and guide me how to improve my mistakes.	101	2.32	1.122
18. All of my teachers ask for computerized assignments.	101	2.43	.898
25. All of my teachers show me the answer sheets of my quizzes, midterm exams and assignments.	101	2.34	.962
28. All of my teachers' give me verbal/written feedback on my quizzes, midterm exams and assignments.	101	2.41	1.106
29. I prefer to know only the grades of my quizzes, midterm exams and assignments instead of my teachers' comments.	101	2.34	1.070
33. I expect my teachers to pass me or give me good marks even if my answers are not very good.	101	2.35	1.004

As per the above statistics, most of the learners disagreed that they got good grades for copying the content of assignments. Similarly, most of them disagreed when asked if their teachers told them about mistakes and guided them on how to improve them. Neither did their teachers ask them for computerised assignments nor show them their answer sheets after quizzes, tests, and mid-terms to overcome their mistakes. The most important point they revealed was that they did not receive any verbal or written feedback, though their teachers' comments mattered more than their grades. Most of them also disagreed with the fact that they expected their teachers to pass them even if their answers were not so good.

Table 1.6 Items that the respondents remained neutral about

Items	N	Mean	SD
6. All of our teachers tell us about the number of words our assignments should have.	101	3.08	.997
12. All of my teachers check our assignments very carefully.	101	2.93	1.061
13. In all my assignments, I have to mention the sources that I quote.	101	3.33	1.201
15. I fail my assignments if I copy the content of my assignments from the internet or a book.	101	2.80	.980
16. I always include references at the end of my assignments.	101	3.02	1.10

			4
17. I have to submit at least one research assignment for every course in each semester.	101	3.19	1.14 6
20. I only study the topics of my courses that I expect in quizzes, midterm and final exams.	101	3.04	1.04 8
21. I memorize the content for tests and exams to get good marks.	101	3.25	1.13 5
22. I have discussion with my teachers to know how to overcome my mistakes in tests.	100	3.05	1.09 5
23. I try to discuss difficult points related to my course with my teachers during their office hours.	101	2.78	1.08 3
27. I can contact all my teachers for feedback	101	2.97	1.07 2
31. I can easily discuss my graded quizzes, midterm exams and assignments with all of my teachers.	101	2.93	1.02 2

The above statistics reveal that most of the learners remained neutral when asked whether their teachers told them regarding the number of words in the assignment and whether they checked their assignments carefully. Similarly, they remained neutral about failing the assignments in the case of copying the contents from the internet or books, including references at the end of their assignments and the submission of at least one research assignment for each course. Likewise, most of them neither agreed nor disagreed with the statements that they studied topics that were expected in quizzes, mid-term and final-term exams and memorised the contents for exams and tests. In addition, they were neutral about the discussion of topics with their course teachers to overcome mistakes, discussions with the teachers in their office hours, easily contacting their teachers for feedback, and easily discussing their graded quizzes and mid-term exams with their teachers.

1. Discussion

The researcher observed the existence of both positive and negative washback on teaching and learning. Keeping in view the 15th hypothesis of Alderson and Wall (1993), that "tests will have a washback effect for some learners and some teachers, but not for others" the researcher assumes that the same assessment system can have one type of washback for students and a different type of washback for teachers.

5.1 Positive Washback Reported

The assessment system seems to have a positive impact on students' learning skills because it develops complementary skills in the learners, such as teamwork and confidence. It involves students in problem-solving discussions with their peers, as well as developing their communication skills, critical thinking, and motivation to learn new material. As Wiggins (1998) highlights, "understanding is best revealed when students are faced with new application situations" (p. 116). Similarly, it encourages meaningful and deep learning in the learners. The students study several sources as well as several pages of a single source for completing their assignments, which means they avoid surface learning of the contents. This finding confirms the 10th hypothesis of Alderson and Wall (1993) that states that tests tend to "influence the degree and depth of learning" (p. 120). Further, such a self-studying and problem-solving system provides them with the opportunity to enhance their cognitive abilities and avoid rote-learning. To support this discussion, some of the respondents' comments are mentioned as:

I memorise some points necessary to be memorized, like dates, epochs, the name of the author and their writings, and the critical commentaries made on them by the critics. But do not try to cram the whole of the content. And I think that cramming should be discouraged.

Similarly, two other students expressed their views in the following words;

I don't memorise content for tests and examinations. I just go through different sources about the course content and pick the relevant idea.

Most of the time, our quizzes, midterms and final exams discourage memorization. I myself memorize content for exam like some dialogues from a drama or some quotations etc. But most of the time there is no space for memorization. As students of English literature have to present arguments and to show their critical thinking in paper.

Likewise, another positive impact on their learning techniques that the respondents reported was that, apart from the timely completion of each task, which prepared them for modern-day real-life situations, it enabled the learners to focus and concentrate on the issue under discussion as well as convey their thoughts and expressions within the required limit. In addition, they felt enabled to filter out what was important and organise their ideas in a coherent manner. Furthermore, it developed their ability to integrate the information studied from several sources into their assignments while observing the rules of academic writing as well. Moreover, the assignments were such that they strengthened students' writing skills by asking them to write essays at length.

Likewise, the participants reported that the assessment methods developed their self-directedness and autonomy as learners. It did not make the students copy someone else's work, but instead it motivated them to produce their own work by using their cognitive and creative skills. This 'ownership' and self-regulation enabled the learners to deep process the materials at hand. They searched, evaluated, and mentioned different sources they studied for their assignments. Further, they preferred such quizzes, tests and exams that required their application ideas, experience and knowledge. Furthermore, their attitude towards the assessment system was also positive. The assignments they reported were a means of improvement for them, and they preferred teachers' comments over grades. In the same manner, another instance of positive washback was their increased awareness of self-assessment. Von Elek (1985) highlights that there is a direct relationship between self-assessment, autonomous learning, and positive washback. He assumes that "it enables them to diagnose their weak areas and to get a realistic view of their overall ability and their skills profile" (Elek, 1985,p. 60).

5.2 Negative Washback Reported

The most evident negative washback in the context of this study is on teaching methodology. Timely feedback is not given to learners. The teachers do not inform students about their mistakes. The students are deprived of any verbal or written comments. Similarly, they do not show and discuss the answer sheets of students, and hence, students don't get the opportunity to improve their learning. To supplement this discussion, some of the respondents' comments are given below. When asked about the quality of feedback by their teachers, one of the students commented that:

Institutionally speaking, in the context of..., the culture of "feedback-based learning" is rare to see. Most of the time, I receive my grades but not the feedback. The problem is not only with me, but in fact, with everyone in my class. I just cannot say anything about the quality due to my inexperience.

Yet another one expressed his views in the following words;

Most of our teachers don't give us any feedback at all. Some of them even don't show us our papers and assignments. We only found out about our grades when the results were displayed. Hardly one or two teachers discuss the paper with us or our grades; otherwise, most of the teachers don't give us feedback. Even when we demand feedback, they neglect our demand.

Likewise, one of them opined;

Very rarely do teachers give feedback to students individually. They mostly appreciate good students and do not focus on the other students who get poor grades, in order to discuss and help them overcome their mistakes.

In the same manner, most of them expressed dissatisfaction with teachers' feedback. This indicates poor student-teacher interaction, which not only sets the trajectory for students' academic careers from as early as the preschool years (Mashburn et al., 2008), but also develops students' motivation to learn (Harter, 2012).

Secondly, the participants reported that teachers did not instruct the students clearly about the topics of the assignments, for which they had to study several sources just to understand the meaning of the assignment topics. This shows that there is negative washback perceived due to the design of the assessment tasks. "Teachers need to provide students with steps and clear instructions that indicate to students what they have to do and how they have to carry out the task" (Muñoz & Álvarez, 2010). Thirdly, the respondents reported that teachers did not use any advanced teaching methodologies within the modern education system to train the students regarding the use of technology for the betterment of their performance. Therefore, the students did not get the opportunity to submit computerized assignments and develop computer skills. This implies that there is an imbalance between learning goals and the assessment system. Assessment by teachers is not a continuous process to improve learning. The teachers seem to be unaware of the importance of the washback effect on their teaching styles, which is leading to students' demotivation of the language.

Conclusion and Recommendations

The findings of the study confirmed the strong washback of the undergraduate assessment system on students' learning strategies and teachers' methodology. The analysis revealed that the assessment system had a positive washback on learners because it developed students' complementary skills along with language skills, motivation for deep-learning, self-directedness in using English, and self-assessment in them. However, the assessment methods under investigation had a negative washback on teachers' methodology. The teachers were preoccupied with curriculum coverage, and did not use assessment for learning. The learners did not receive detailed and timely feedback. Further, they were not given instruction about assessment tasks or any opportunity to develop their computer skills.

REFERENCES

Aftab, A., Qureshi, S., & William, I. (2014). Investigating the washback effect of the Pakistani

- Intermediate English Examination. *International Journal of English and Literature*, 5(7), 149-154.
- Ahmad, S., & Rao, C. (2012). Examination washback effect: Syllabus, teaching methodology and the learners' communicative competence. *Journal of Education and Practice*, 3(15), 173-183.
- Alderson, J. C., & Wall, D. (1992) Does washback exist? "The Educational and Social Impacts of Language Tests", *Language Testing Research Colloquium*.
- Alderson, J. C., & Wall, D. (1993). Does washback exist? *Applied linguistics*, 14(2), 115-129.
- Alderson, J.C. (2004) *Washback in Language Testing: Research Context and Methods*, ed. by L. Cheng, Y. Watanabe, A. Curtis (Eds), Lawrence Erlbaum Associates, Inc, Mahwah.
- Bachman, L. F., Palmer, A. S. (1996) *Language testing in practice*. Oxford, England: Oxford University Press.
- Bailey, K. M. (1996) Working for washback a review of the washback concept in language testing. *Language Testing* 13(3), 257-279.
- Biggs, J.B. (1995) Assumptions Underlying New Approaches to Educational Assessment. *Curriculum Forum*. 4(2), 1-22.
- Bogdan, R. & Biklen, S. K. (1998). *Qualitative Research for Education: An Introduction to Theory and Methods*. London: Allyn & Bacon.
- Cheng, L. (1997). How does washback influence teaching? Implications for Hong Kong. *Language and Education*, 11(1), 38-54.
- Cheng, L. (1998). Impact of a public English examination change on students' perceptions and attitudes toward their English learning. *Studies in Educational Evaluation*, 24(3), 279-301.
- Cheng, L. (1999) Changing assessment: Washback on teacher perspectives and actions. *Teaching and Teacher Education*, 15, 253-271.
- Cheng, L. (2001). Washback studies: Methodological considerations. *Curriculum Forum*, 10(2), 17-32.
- Cheng, L. (2003). Looking at the impact of a public examination change on secondary classroom teaching: A Hong Kong case study. *Journal of Classroom Interaction*, 38(1), 1-10.
- Cheng, L. (2004). The washback effect of a public examination change on teachers' perceptions toward their classroom teaching. In L. Cheng, Y. Watanabe, & A.

- Cheng, Y. Watanabe, A. Curtis (eds.) (2004). Washback in Language Testing: Research Context and Methods. Lawrence Erlbaum Associates, Inc., Mahwah.
- Creswell, J. W. (2014). Research Design: qualitative, quantitative and mixed method approaches. 4th Ed. United States of America: SAGE Publications, Inc.
- Cronbach, L. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16, 297-334.
- Curtis (Eds.), Washback in language testing: Research contexts and methods. Mahwah, NJ: Lawrence Erlbaum Associates. 146-170.
- Davies, A. (1997) Demands of being professional in language testing. *Language Testing*, 14, 328-339.
- Denscombe, M. (2017). EBOOK: The Good Research Guide: For Small-Scale Social Research Projects. McGraw-Hill Education (UK).
- Ferman, I. (2004) The washback of an EFL national oral matriculation test to teaching and learning. Washback in Language Testing: Research Context and Methods, ed. by L. Cheng, Y. Watanabe, A. Curtis (Eds). Lawrence Erlbaum Associates, Inc, Mahwah. 191-210.
- Green, A. (2007). IELTS washback in context: Preparation for academic writing in higher education (Vol. 25). Cambridge University Press.
- Hafsa, N. E. (2019). Mixed methods research: An overview for beginners researchers. *Journal of Literature, Language and Linguistics*, 58, 45-49.
- Harmer, J. (2007) The Practice of English Language Teaching (4th Ed). Cambridge, England: Pearson Longman.
- Harter, S. (2012). Emerging self-processes during childhood and adolescence. In M. R. Leary & J. P. Tangney (Eds.), *Handbook of self and identity* (pp. 680-715). The Guilford Press.
- Hughes, A. (1989) Testing for language teachers. Cambridge, England: Cambridge University Press.
- Hughes, A. (2003) Testing for language teachers (2nd ed.). Cambridge, England: Cambridge University Press.
- Likert, R. (1931) A technique for the measurement of attitudes: *Archives of Psychology*. New York: Columbia University Press.
- Mashburn, A. J., Pianta, R. C., Hamre, B. K., Downer, J. T., Barbarin, O. A. Bryant, D., ... &

- Howes, C. (2008). Measures of classroom quality in prekindergarten and children's development of academic, language and social skills. *Child development*, 79 (3), 732-749.
- Messick, S. (1996) Validity and washback in language testing. *Language Testing*, 13(3), 241-256.
- Morrow, K. (1986) The evaluation of tests of communicative performance. In M. Portal (Ed.), *Innovations in language testing*. London, England: NFER/Nelson. 1-3.
- Muñoz, A. P., &Álvarez, M. E. (2010). Washback of an oral assessment system in the EFL classroom. *Language testing*, 27(1), 33-49.
- Newfields, T. (2005) Voices in the Field: An Interview with Yoshinori Watanabe Voices in the Field: An Interview with Yoshinori Watanabe. *JALT Testing & Evaluation SIG Newsletter*. 9 (1). 5-7
- O'malley, J., & Price, M. (1996). *Authentic Assessment for English Language Learners*. Pearson Education ESL
- Pimentel, J. (2010). A note on the usage of Likert Scaling for research data analysis. *USM R & D journal*, 18, 109-112.
- Popham, W. J. (1987) The merits of measurement-driven instruction. *ThePhi Delta Kappan*, 68, 679-682.
- Quansah, F. (2017) The Use of Cronbach Alpha Reliability Estimate In Research Among Students In Public Universities In Ghana. *African Journal of Teacher Education*, 6.
- Ren, Y. (2011) A study of the washback effects of the College English Test (band 4) on teaching and learning English at tertiary level in China, *International Journal of Pedagogies and Learning*, 6(3). 243-259.
- Resnick, L. B., & Resnick, D. P. (1992) Assessing the thinking curriculum: New tools for educational reform. B. R. Gifford & M. C. O'Connor (Eds.), *Changing assessments: Alternative views of aptitude, achievement and instruction* London: Kluwer Academic. 37-75.
- Robinowitz. (2010). From principles to practice: An embedded assessment system. *Applied Measurement in Education*, 13(2), 181-208.
- Shamim, F. (2006) *Case Studies of Organization of English Language Teaching in Public Sector*

- assessment tasks, learning strategies and learning materials. *Kashmir Journal of Language Research*, 16(2), 109.
- van Lier, L. (1989). Reeling, writing, drawling, stretching and feinting in coils: oral proficiency interviews as conversation. *TESOL Quaterly* 23, 489-508.
- von Elek, T. (1985). 1985: A test of Swedish as a second language: an experiment in self-assessment. In Lee, YP, Fok, CYY, Lord, R. and Low, G., editors, *New dimensions in language testing*, Oxford: Pergamon Press, 47-58.
- Weir, C. J. (1990) *Communicative Language Testing*. New York. Prentice Hall.
- Wiggins, G. (1998). *Educative Assessment. Designing Assessments To Inform and Improve Student Performance*. Jossey-Bass Publishers, 350 Sansome Street, San Francisco, CA 94104.
- Xu, Q. Liu, J. (2018) *A Study on the Washback Effects of the Test for English Majors (TEM)*. Singapore: Springer Nature Singapore.