E-assessment: Wash-Back Effects and Challenges
(Examining Students’ and Teachers’ Attitudes Towards E-tests)

Ibtessam Abdulaziz Binnahedh
Department of Curricula and Teaching Methods, College of Education, Shaqraa University, KSA

Abstract—Online testing that includes multiple-choice tests and assessment of problem-solving skills is considered to be the main form of e-assessment. In comparing e-tests to paper-based tests, it is found that e-tests are more accurate and faster than the traditional method. Many obstacles represent challenges for applying electronic tests. Both teachers and students are affected by these challenges. The current paper examines both students’ and teachers’ perceptions towards e-tests in addition to students’ wash-back effects of e-tests. It also reviews and discusses obstacles and challenges that may not only affect students and teachers but also may result from the institution, infrastructure, and curriculum. The sample of the study includes a group of secondary school students (n=75) enrolled in the schools of Al-Dawadmi directorate, KSA, and a sample of EFL teachers (N=41) working in the same governorate. Instruments include students’ perception questionnaires and teachers’ perceptions questionnaires to examine their attitudes towards e-tests. Results of the questionnaires were analyzed statistically using the SPSS program. Findings indicated moderate washback effects and attitudes towards e-tests students’ perspectives as well as a high level of students’ perceptions towards e-tests. But teachers’ perceptions proved to be below average.

Index Terms—Wash-back effect, Negative wash-back, E-tests, electronic assessment

I. INTRODUCTION

In EFL classes, students show a convenient level and participation represented in their engagement with the teacher. But when examined, the level shown indicates a different predictor for success. So, in examining and searching for reasons that hinder getting high scores in exams, studies showed many effects of exams whether negative or positive. Most factors are psychological and others are logical. Effects of having exams are called washback effects. Hence, exams have different effects on both teachers and learners. The washback effect is defined as "the influence of testing on teaching and learning (Alderson & Wall, 1993)." The extent to which the implementation and usage of a test encourage language teachers and learners to do things they would not otherwise do that enhance or restrict language learning," writes Messick (1996, p. 241). According to Wall (1997, p.11), test impact refers to any of a test's consequences on persons, policies, or practises in the classroom, school, learning environment, and community as a whole. Tests can have both adverse effects, as well as wash-back. The term "positive wash-back" refers to the expected results of a test. A test, for example, may motivate learners to study more or establish a link between norms and education. Negative wash-back refers to a test's unintended and detrimental consequences. For instance, the training may place an excessive emphasis on test preparation at the expense of other activities. Individual teachers and students, as well as entire courses and programmes, might be involved in test wash-back. E-testing is a fast increasing area of e-assessment that entails delivering tests and assessments on a computer display, either utilising local or web-based technology. The current paper will examine and address obstacles and problems that may arise from the organization, facilities, and curriculum, as well as from students and teachers. The paper also discusses the wash-back impact of e-tests from the viewpoints of both students and instructors.

According to Cheng and Watanabe (2008), there are at least two major types of washback studies: those relating to traditional, objective tests, which are perceived to have had primarily negative effects on the quality of teaching and learning, and those relating to other types or areas of washback or backwash studies. The second sort of research, on the other hand, has found no impact on teaching and learning. Moreover, many of these researches have shifted their focus to figuring out how washback or backwash is employed to alter teaching and learning. The concept is rooted in the notion that tests or examinations can and should drive teaching, and hence learning.

II. BACKGROUND

Tests, especially large-scale, high-stakes ones, tend to exert more influence on their stakeholders, such as learners, teachers, and school administrators (Zou & Xu, 2017). In secondary school, online assessments are becoming increasingly important. The nature of a test can have an immediate impact on the participants' expectations and attitudes toward their teaching and learning tasks. These expectations and attitudes, in turn, can influence what participants do in
the course of their work (process), such as practising the types of things that will appear on the test, affecting learning outcomes.

Cheng and Curtis (2004) agree that who performs the investigation in a specific situation determines whether the results of tests are positive or negative. They also say that it is important to consider where the investigation takes place — school or university settings; were — the period and length of utilizing those evaluation practices; and how the investigation is conducted. Why — the rationale; and how — the different approaches used by different participants within the context (p. 8)

III. STUDIES

Han's paper (2021) examines the backwash of reformed English, drawing on theoretical models and empirical evidence from both at home and abroad. CET-6 used a questionnaire to assess learners' listening skills. The paper was surveyed using quantitative analysis techniques with 60 samples in many public universities to conduct the study. Following data collection and analysis, the authors concluded that this test has a substantial washback impact on student learning.

Ali and Hamid (2020) investigated the factors that lead to a negative washback impact on English language teaching (ELT) in Bangladeshi secondary schools. It was suggested in the study that maintaining a testing-teaching causal relationship may be too simplistic since testing and teaching are intertwined, a complex of social psychological, political, economic, and data-driven accountability variables that influence language instruction. In cases where language evaluations are assumed to impede language instruction, these findings have implications for language testing and evaluation researchers, policymakers, programme designers, and administrators.

In a Turkish EAP background, Saglam (2018) published a mixed-method study that looked at the washback effect from a locally generated, theme-based, high-stakes English language proficiency test in tertiary education. The aim was to see how much washback on teaching could be caused by an interactive theme-based English proficiency test that was designed to represent authentic language usage in Turkey's tertiary education system. Classroom observations and focus group interviews with 14 teachers from the Preparatory English Language Program were used to collect data (PEP). Observations in the classroom were also carried out. The results revealed that test effects on teaching were both positive and negative. The findings have suggested that both material and teaching methodology are affected.

The study by Hungs (2012) examined the positive and negative washback effects of e-portfolio tests on learning. This evaluation project included eighteen English as a Foreign Language (EFL) student teachers enrolled in a Master's program in Teaching English to Speakers of Other Languages. Multiple instruments were used to collect data, including interviews, observations, text analysis, and reflective journals. The results indicate that e-portfolio evaluations have positive learning washback effects, such as creating a community of practice, promoting peer learning, and improving content awareness learning. E-portfolio evaluations, on the other hand, have certain undesirable side effects, such as learning anxiety caused by wider audiences and hostility to technology.

Allen (2016) examined the IELTS Academic exam's consequential validity, concentrating on the impact of washback on learners' test preparation strategies and score gains, as well as the mediating factors affecting washback when learners in an EFL setting are not enrolled in test preparation courses. The results of the tests showed that the IELTS Test had a positive impact on learners' language abilities and test preparation strategies, especially in terms of productive skills, which learners had previously ignored in their language studies. However, to ensure positive washback in EFL situations and the absence of guidance, many mediating factors must be discussed.

Qi (2005) found that testing procedure (the format of tests) can influence the material of student learning regardless of the test designers' intentions, and that key stakeholders' goals, as well as test stakes, play a key role in the body of student learning.

Tayeb et al (2014) investigated the washback effect of the General Secondary English Examination in Yemen. They wanted to get a better understanding of the relationship between the washback effect and teaching and learning aspects. The study focused on eight pedagogical dimensions: four of them concerned the teachers (teaching methods, teaching experiences, content assessment, and beliefs) and four concerned the students (learning styles, learning activities, attitudes, and motivation). Instruments used included a questionnaire and a semi-structured interview. According to the findings, the test had a significant impact on learners and teachers, particularly on teaching techniques and learning styles. The study shows that the exam has a washback effect on the components of Yemen's language teaching and learning processes.

IV. PROBLEM

COVID-19 pandemic resulted in serious problems and critical situations in all fields of life since its emergence in 2020. The field of education is affected greatly whether in giving regular sessions or in exams. As for school teaching periods, the experiment was successful to a great extent when applying the online- sessions through the platform of Madrasty and Research Teams. Still, secondary school students, as well as teachers, face the difficulty of applying the e-tests during formative and summative exams. The current research examines and uncovers the negative and positive
washback effects, challenges, and attitudes towards e-tests among both teachers and students. That is to evaluate the situation and suggest solutions for more benefits and improvements.

V. AIMS OF RESEARCH

The current research aims at the following:
1. Examining secondary school students towards e-tests. Examining secondary school students washback effects of e-tests.
2. Examining teachers' attitudes towards e-tests.
3. Uncovering challenges and promises that face the e-tests through reviewing the literature concerning e-learning and e-assessment.

VI. QUESTIONS OF THE STUDY

1. What are secondary school students' attitudes and perceptions towards e-tests?
2. What are the washback effects of e-tests among secondary school students?
3. What are teachers' attitudes and perceptions towards e-tests?
4. What are the challenges and promises of e-tests during the phase of the COVID-19 pandemic?

VII. HYPOTHESES

a. There are positive attitudes towards e-tests among students of Dawadmi secondary schools
b. There are positive attitudes towards e-tests among teachers of Dawadmi secondary schools
c. There are positive washback effects of e-tests among Dawadmi secondary schools students

VIII. METHOD

Participants
1. Female secondary school students (n= 47) in Al-Dawadmi city, Riyadh Province, KSA
2. Female secondary school teachers (n=9) in Al-Dawadmi city, Riyadh Province, KSA

Instruments
1. Teachers' Perceptions and Attitudes Survey (TPAS).
   The survey was adopted from Da'asin (2016) and adapted to meet the data required in the current study. The TPAS was presented to several jury members (n=6) who specialized in teaching English as a foreign language and had experience in teaching online through the COVID-19 pandemic. They were asked to modify and validate the survey.
   For the sake of measuring reliability, the survey was applied to a group of teachers (n=5) different from the study sample. Responses were analyzed using SPSS. Alpha Cronbach's reliability was (0.831) which means it has high reliability.
2. Students' Perceptions, and washback effects Survey (SPACES).
   The survey was adopted from Meseke et al (2010) and Tayeb et al (2014) and was adapted and modified to match the data required in the present study.
   The SPACES was presented to several jury members (n=6) who specialized in teaching English as a foreign language and had experience in teaching online through the COVID-19 pandemic. They were asked to modify and validate the survey.
   For the sake of measuring reliability, the survey was applied to a group of students (n=30) different from the study sample. The alpha Cronbach reliability was measured using the SPSS program. It was (0.781) which means acceptable reliability.

IX. RESULTS

Question 1:
1. What are secondary school students' attitudes and perceptions towards e-tests?
Table (1) shows descriptive statistics for Students’ Perceptions, Attitudes towards Online Testing, from which we find that the highest average was awarded to the items receiving ranks from 1 to 7. Items receiving ranks from 8 to 9 were awarded a moderate level with means; 03.09, 2.93, and Std. 1.11 and 1.16. The weighted average of this section entitled: Students’ Perceptions, Attitudes towards Online Testing, was 3.40 with Std. D. 0.62, which is regarded as high; since the intervals of levels are as follows:

- **Low level**: [1-2.59]
- **Moderate**: [2.60-3.39]
- **High level**: [3.40-5]

**Question 2:**

What are the washback effects of e-tests among secondary school students?
Table 2 shows descriptive statistics for Students’ washback effects towards Online Testing. We can see that all items received a moderate average with a general weighted mean of 3.06. It is considered as a moderate average.

**Question 3**
What are teachers’ attitudes and perceptions towards e-tests?
Table (3) shows descriptive statistics for "teachers' attitudes and perceptions towards e-tests". The moderate averages were awarded to the items which occupied ranks from 1 to 5. While a Low average was awarded to items that occupied ranks from 6 to 12. The weighted means of teachers' attitudes and perceptions towards e-tests was 2.52 with Std. 1.33, which means low level.

| Items                                                                 | Frequency | S. disagree | Disagree | Neutral | Agree | S. Agree | M.          | SD       | Rank | degree |
|----------------------------------------------------------------------|-----------|-------------|----------|---------|-------|----------|-------------|----------|-------|--------|
| I think it is appropriate to allow students to try before they finally attend the E-exam. | N 16 | 39.02 | 2 | 4.87 | 4 | 9.75 | 11 | 26.82 | 8 | 19.51 | 2.82 | 1.64 | 1 | M |
| Drafting the paragraphs of the E-exam in the form of multiple-choice questions is appropriate. | N 15 | 36.58 | 7 | 17.07 | 5 | 12.19 | 9 | 21.95 | 4 | 9.75 | 2.46 | 1.45 | 10 | L |
| The number of electronic exam questions is sufficient. | N 19 | 46.34 | 3 | 7.31 | 4 | 9.75 | 10 | 25.39 | 5 | 12.19 | 2.48 | 1.56 | 9 | L |
| The electronic exam system is clear and specific. | N 16 | 39.02 | 3 | 7.31 | 7 | 17.07 | 11 | 26.82 | 4 | 9.75 | 2.60 | 1.48 | 5 | M |
| Students do not feel worried when using the computer. | N 16 | 39.02 | 6 | 14.63 | 2 | 4.87 | 12 | 29.26 | 5 | 12.19 | 2.60 | 1.54 | 5 | M |
| E-exam helps extract results quickly. | N 17 | 41.46 | 2 | 4.87 | 24.87 | 11 | 26.82 | 9 | 21.95 | 2.82 | 1.70 | 1 | M |
| E-exam is serious. | N 15 | 36.58 | 6 | 14.63 | 6 | 14.63 | 10 | 25.39 | 4 | 9.75 | 2.56 | 1.44 | 6 | L |
| E-exam regulations are clear and easy to understand. | N 16 | 39.02 | 0 | 5 | 12.19 | 1843.9 | 2 | 4.87 | 2.75 | 1.47 | 2 | M |
| E-exam times are suitable for students. | N 17 | 41.46 | 4 | 9.75 | 5 | 12.19 | 13 | 31.7 | 2 | 4.87 | 2.48 | 1.43 | 9 | L |
| E-exam results are reliable. | N 16 | 39.02 | 8 | 19.51 | 4 | 9.75 | 12 | 1 | 2.43 | 2.56 | 1.33 | 13 | L |
| E-exam does not affect students’ success in the comprehensive exam. | N 17 | 41.46 | 6 | 14.63 | 5 | 12.19 | 13 | 31.7 | 0 | 2.34 | 2.34 | 1.31 | 15 | L |
| E-exam measures what it is intended to measure(valid). | N 17 | 17.07 | 6 | 14.63 | 5 | 12.19 | 13 | 31.7 | 0 | 2.51 | 2.51 | 1.45 | 8 | L |
| E-exam is effective. | N 17 | 17.07 | 3 | 7.31 | 7 | 17.07 | 11 | 26.82 | 3 | 7.31 | 2.51 | 1.48 | 8 | L |
| Students do not need external help when using the computer. | N 18 | 43.9 | 3 | 7.31 | 3 | 7.31 | 15 | 36.58 | 2 | 4.87 | 2.41 | 1.46 | 11 | L |
| E-exam enjoys consensus. | N 19 | 46.34 | 3 | 7.31 | 4 | 9.75 | 13 | 31.7 | 2 | 4.87 | 2.53 | 1.38 | 7 | L |
| E-exam helps control the quality of the comprehensive exam. | N 16 | 39.02 | 2 | 4.87 | 10 | 25.39 | 1126.8 | 2 | 4.87 | 2.51 | 1.46 | 8 | L |
| E-exam limits cheating attempts. | N 17 | 17.07 | 5 | 12.19 | 2 | 4.87 | 15 | 36.58 | 2 | 4.87 | 2.07 | 1.33 | 16 | L |
| The distribution of scores on E-exam papers is fair. | N 23 | 56.09 | 2 | 4.87 | 7 | 17.07 | 8 | 36.58 | 1 | 2.43 | 2.65 | 1.55 | 4 | M |
| The philosophy of shifting towards an E-exam is justified. | N 17 | 17.07 | 3 | 7.31 | 2 | 4.87 | 15 | 36.58 | 4 | 9.75 | 2.70 | 1.47 | 3 | M |
| The exam time is enough to answer all questions. | N 15 | 36.58 | 3 | 7.31 | 5 | 12.19 | 15 | 36.58 | 3 | 7.31 | 2.48 | 1.45 | 9 | L |
| E-exam does not raise the level of anxiety and stress. | N 17 | 17.07 | 5 | 12.19 | 3 | 7.31 | 14 | 3.5 | 2 | 4.87 | 2.39 | 1.46 | 12 | L |
| E-exam helps raise the efficiency of student achievement. | N 18 | 43.9 | 6 | 14.63 | 3 | 7.31 | 11 | 26.82 | 3 | 7.31 | 2.46 | 1.39 | 10 | L |

Weighted mean | 2.52
Std. Deviation | 1.33
As learners and teachers are the principal focus of washback studies (Johnson & Shaw, 2019), the current research examines students’ perceptions and washback effects of e-tests towards e-tests as well as the teachers’ perceptions towards e-tests. Figure 1 shows the means of students’ Perceptions, Attitudes towards Online Testing. The general and weighted mean was high. This result answers the first question of the study.

Figure 1. Students’ Perceptions, Attitudes towards Online Testing

Figure 2 shows the means of washback effects of e-tests from students' perspectives. The weighted mean was 3.06 which proved to be moderate.

Figure 2. Students’ washback effects towards Online Testing

Figure 3. shows the teachers’ attitudes and perceptions towards e-tests with a weighted mean of 2.52 which proved to be below.
The results of the current study agree with Jiamin et al (2021) and Saglam (2018). Although around 26% of students expressed their suffering from internet connectivity, 57% felt confident with online testing and 57% enjoyed using computers and e-learning environments in testing and about 62% of views enjoyed and recommended using e-learning environments in testing language.

Considering washback effects of e-tests, we find that 8% of views disagreed with the point that e-tests lead teachers to use a teaching-to-the-test approach in the class which is regarded as a negative washback effect. While 29.33% of responses were neutral and only 26.66% agreed. This means that the effect of e-tests here is positive, not negative. In addition, 24% of students view that e-tests motivate teachers to improve their methodology in teaching English. In terms of other behaviors of teachers that are affected by the e-tests, about 45.33% of students agreed that teachers tend to teach only the points similar to that of the exam. As for students, they expressed their fear and stress of online tests with about 32%. This means that e-tests affect both students and teachers’ behaviors even if they have no problems in using e-learning environments or platforms.

Teachers showed varied views towards e-tests. The highest rank was given to the items (I think it is appropriate to use e-learning environments in testing and about 62% of views enjoyed and recommended using e-learning environments in testing language.

Good washback effects will be achieved, according to Cheng and Curtis (2004), when teachers and learners have a positive mindset toward testing and engage proactively and cooperatively toward evaluation and learning goals.

According to Hughes’ (1989) definition, the washback effect has been used to refer to the way a test affects teaching materials and classroom management. Negative washback effects refer to undesired impacts on teaching and learning in the traditional testing setting, such as an increasing focus on memory, rehearsing exam strategies rather than language learning tasks, excessive test anxiety, and inability to build broad knowledge (Curtis, 2004). Washback has also been formulated as the influence of tests which lead “teachers and learners to do things they would not necessarily otherwise do” (Alderson & Wall, 1993, p.17). Positive washback occurs when a testing procedure brings about good teaching practice (Taylor, 2005) and more focused teaching (Johnson & Shaw, 2019).

Hughes’ (1994) washback trichotomy model divides test effects into three categories: participants, procedure, and product. Participants, for starters, refer to stakeholders like students, teachers, administrators, materials writers, and publishers, whose perspectives of the teaching and learning process may be influenced by exams. Switching exams from paper-based to electronic tests affected the participants’ behaviors to achieve the highest outcomes. Teachers were affected greatly as they tried to change methods to decrease the gap between instruction and exams. About 30% of responses agreed that the teacher used teaching to the test strategy, while more than 55% of responses agreed that teachers explained grammatical rules directly the be familiar with the e-tests MCQs strategies. This process led students to ask continuously about tactics that facilitate answering multiple choices questions (about 49% of responses). As for the psychological impact of tests generally and of e-tests specifically, there was a prominent low impact of the later one. A low percentage of responses (38%) agreed that e-tests cause anxiety and worry on them.

The practice that relates to initiatives in teaching and learning such as materials development, syllabi design, instructional adjustments, and method is the second sort of washback effect. Implementing e-tests affects students’ creativity because of the nature of questions. All questions are based on MCQs which force students not to train themselves on creative and high-level answers. So this creates a gap between curriculum and creativity. According to...
Johnson and Shaw (2019), prioritising test achievement over knowledge and comprehension has negative effects for instructors; that is, teachers believe exams are insufficient and that there is a disconnect between what is tested and what is taught. Answers in the exam make students similar in abilities except for essay questions. Teachers are motivated to use different methods but they were not advanced. The methods used were to meet the demands of the e-tests.

Finally, the product considers students' intake, competencies, and learning quality (Bailey, 1996, p. 262). Tests have two effects: they impact participants (teaching staff, students, and materials writers involved in test preparation, as well as the perceptions and attitudes they bring to the assignment), and they cause them to change their processes (teaching and learning behaviours). As a result, these have an impact on learning outcomes (Green, 2007a, p.78). Since the nature of e-tests focuses on some low and intermediate levels and skills ignoring sometimes high levels and creative skills, the product is affected; some skills are developed while others are neglected.

ACKNOWLEDGEMENTS

The author would like to thank the Deanship of Scientific Research at Shaqra University, Saudi Arabia for supporting this work.

REFERENCES

[1] Alderson, J. C., & Wall, D. (1993). Does washback exist? *Applied Linguistics, 14*, 115–129.
[2] Ali, M. M., & Hamid, M. O. (2020). Teaching English to the test: Why Does negative washback exist within secondary education in Bangladesh? *Language Assessment Quarterly, 17*(2), 129-146.
[3] Cheng, L., & Curtis, A. (2004). Washback or backwash: A review of the impact of testing on teaching and learning. In L. Cheng, Y. Watanabe, & A. Curtis (Eds.), *Washback in language testing: Research contexts and methods* (pp. 3–18). Mahwah, NJ: Lawrence Erlbaum Associates.
[4] Da'asin, K. A. (2016). The attitude of Ash-Shobak University College Students to E-Exam for Intermediate University Degree in Jordan. *Journal of Education and Practice, 7*(9), 10-17.
[5] Hughes, A. (1994). *Backwash and TOEFL 2000*. Unpublished manuscript, commissioned by Educational Testing Service (ETS). The University of Reading.
[6] Hung, S. T. A. (2012). A washback study on e-portfolio assessment in an English as a Foreign Language teacher preparation program. *Computer Assisted Language Learning, 25*(1), 21-36.
[7] Jiamin, X., Jinyan, L., & Tianyi, M. (2021). The wash-back effect of reformed CET 6 listening comprehension test. *Asian Journal of Education and Training, 7*(1), 70-73.
[8] Johnson, M.& Shaw, S.(2019). What is computer-based testing washback, how can it be evaluated and how can this support practitioner research?, *Journal of Further and Higher Education, 43*(9), 1255-1270, DOI: 10.1080/0309877X.2018.1471127
[9] Meseke, Christopher A., Rita Nafziger, and Jamie K. Meseke. (2010). "Student attitudes, satisfaction, and learning in a collaborative testing environment." *Journal of Chiropractic Education 24*, no. 1: 19-29.
[10] Messick, S. (1996). Validity and wash-back in language testing. *Language testing, 13*(3), 241-256.
[11] Saglam, A. L. G. (2018). Can exams change how and what teachers teach? Investigating the washback effect of a university English language proficiency test in the Turkish context. *Eurasian Journal of Applied Linguistics, 4*(2), 155-176.
[12] Tayeb, Y. A., Abd Aziz, M. S., Ismail, K., & Khan, A. B. M. A. (2014). The wash-back effect of the general secondary English examination (GSEE) on teaching and learning. GEMA Online® Journal of Language Studies, 14(3),83-103.
[13] Wall, D. (1997). Impact and wash-back in language testing. In C. Clapham & D. Corson (Eds.). *Encyclopedia of Language and Education* (pp. 291-302). Dordrecht: Kluwer Academic Publishers

Ibtisam Abdulaziz Binnahedh was born in Riyadh, Saudi Arabia. She got a master degree in Curricula and Teaching Methods from King Saud University, Riyadh, KSA, in 2009. She also got PhD in Curricula and Teaching Methods from King Saud University, Riyadh, KSA, in 2017.

She holds leadership positions in the College of Education, at Shaqraa University. Currently, she holds the position of the Dean of College of Education, Shaqra University. She is interested in researching the recent trends in curricula and teaching methods.

Dr Binnahedh is an active member in many of the committees of preparing programs and plans in Shaqra University such as the strategic plan committee and the developmental initiatives as well as curricula development committee.