
RESEARCH ARTICLE

EFL High-Stakes Exams: Are We Leading Teachers as Language Teachers or Test Teachers?

Mohammad M. A. Homran¹ and Sahail M. M. Asassfeh² ✉

¹*Teacher of English at the Ministry of Education, Jordan*

²*Professor at the Hashemite University, Jordan*

Corresponding Author: Sahail M. M. Asassfeh, **E-mail:** sabash@hu.edu.jo

ABSTRACT

This paper reports on a study investigating the washback effects of the Jordanian General Secondary Certificate English Exam (JGSCEE), a high-stakes exit exam for second secondary school students (12th grade) in Jordan, on EFL Teaching from a teacher's perspective. It also investigates whether teachers' views are influenced by certain variables, such as gender, school type, and years of experience. A questionnaire that covered three dimensions of the teaching process (planning, implementation, and assessment) was distributed face-to-face to a sample of 171-second secondary grade EFL teachers. Data were analyzed using descriptive and referential statistics via SPSS (Version 23). The results showed that the JGSCEE had a great influence on teachers' practices at the level of each of the three teaching process dimensions. As for the classifying variables, the results showed no statistically significant differences ($\alpha= 0.05$) associated with the years of experience. On the other hand, the responses varied significantly according to gender in the planning dimension in favor of female teachers and according to the school type in favor of private ones on the three dimensions.

KEYWORDS

Washback; EFL teaching; EFL assessment; Jordanian General Secondary School Certificate Exam

ARTICLE INFORMATION

ACCEPTED: 02 October 2023

PUBLISHED: 26 October 2023

DOI: 10.32996/jweep.2023.5.3.4

1. Introduction

Educational systems inevitably need assessments, and the value of any assessment is associated with the level of impact assessment results have on stakeholders. Particularly for EFL learners, international standardized tests (e.g., TOEFL and IELTS) are of paramount significance as they might make or break examinees' academic future. National exams as well have a comparable weight, especially when taken as the sole basis for shaping students' university admission.

In fact, the intent, formats, resources, and choices made in response to assessment results vary across contexts and instances. In the case of Jordan as an example and context of interest for this study, the educational system is divided into three stages; preschool, basic school and secondary school (Ministry of Education, 2014). Academic and vocational programs are the two major routes in secondary school. The academic stream includes both a literary stream and a scientific stream. It is widely known that the educational system in Jordan is largely test-driven and where exams, particularly public ones, are given excessive weight (Haddadin, 2006). All Jordanian students must take a national unified exit exam by the time they graduate from secondary school as part of Jordan's centralized university admission procedures in addition to the country's centralized university admission policy. A student's mean score on the General Secondary School Certificate Exam (GSSCE) is the most important criterion that determines a student's university admission. This, in turn, explains why students, teachers, and parents place so much importance on these results. Thus, it is vital to examine the impact of this frequent practice on the educational process, specifically the washback effects of this paper-and-pencil-only exam. The JGSCEE is considered as a test with high stakes. As a result, the effort and considerable funds spent by parents, schools, and the Jordanian Ministry of Education on assisting students in performing well on the test indicate the perception that the test has a very significant impact on society and educational institutions (Shatnawi, 2005). Since

the JGSCEE is the means by which students might achieve their future goals, teachers focus their efforts on presenting the prescribed curriculum while employing commercial books that contain several previous years' tests to provide their students with practice (Ghadi & Al-Jamal, 2008). As a result, humanistic and communicative approaches are dismissed as an unaffordable luxury, and this traps teachers in an endless exam preparation cycle (Prodromrou, 2006). According to this scenario, EFL teachers in Jordan are unofficially obliged to teach to the test, particularly at the secondary levels (Haddadin, 2006). The purpose of this study is to unveil how JGSCEE contributes to shaping EFL teachers' practices in response to the responsibility of developing students' communicative competence in an era of internationalization and globalization characterized by an excessive level of competitiveness and taking some classifying variables into account (teachers' gender, experience, and school type. study).

2. Study Background and Literature Review

Washback has a plethora of definitions in applied linguistics research. For example, it refers to the extent to which a test influences teachers and students to behave in specific ways (Alderson & Wall, 1993). It can also be viewed as a process that affects many aspects of teaching and learning (Ozmen, 2011; Wang, 2010). Messick (1996) defines the washback effect in a wider context as the extent to which the use of tests influences teaching and learning by forcing learners to do things they would not normally do to inhabit or encourage language learning. This influence extends beyond teaching and learning to include teachers' and students' attitudes and behaviors in response to external testing (Cheng, 2005 ; Schohamy, 2020). The test washback effect is not a strange idea in the testing literature; it describes the impact of foreign language exams on teaching and learning. This term did not catch the attention of test researchers until the early 1990s. Prior to that time, researchers used a wide range of terms. The concept of test impact (Bachman & Palmer, 2000; Baker, 1994) refers to the impact of a test on teaching and learning. The term "systematic validity" (Frederickson & Collins, 1989) refers to the process by which assessments are integrated into the educational system.

Many studies have addressed exam consequences on the teaching process. For example, Manjarres (2005) investigated the washback effect of the National English Examination (NEE), a high-stakes test that is administered prior to university admission in Colombian public schools. The main question of the study was whether the English test had any effect on English instruction and if it assessed students' grammatical and linguistic proficiency. Five lessons were observed for data collection, three students and two teachers were interviewed, and 2003-2004 student exams were examined and compared to classroom practices. According to the findings, there is a strong correlation between test scores and teacher performance, and English language teachers modify their lesson plans to better meet the needs of their students. This relationship is also apparent when teachers rely on additional resources. The results also showed negative washback because listening and speaking were not assessed during classroom tests because they were not included in NEE, and teachers lacked knowledge of how to help students develop their communicative competence. These findings are in line with (Hawky, 2004; Ying, 2005; Wall and Alderson, 1993).

The Basic Competency Test (BCT), a public exam in Taiwan that determines whether a student is admitted to a high school, was also the subject of a study by Chen (2002) to determine how English teachers in junior high schools felt about its curriculum design and instruction. A questionnaire was distributed to 178 teachers, and nine focus groups were interviewed to gather data. The result showed that the BCT has an impact on how teachers plan their lessons. Washback may have an impact on how teachers teach students, but it has little impact on what they teach. The test had an effect on instructional practices, according to Chen, because junior high schools all over the nation had recently received new teaching materials.

In Oman, Al Lawati (2002) investigated the washback effects of the Secondary Certificate English exam on teachers and students' practices, as well as the extent to which these practices were influenced by the gender, experience, and academic stream of the teachers and students. A questionnaire with five dimensions, Grammar and vocabulary, language abilities, exam modification regions, time management, and textbook match and mismatch, was given to 54 teachers and 157 students. According to the findings, the Secondary Certificate English exam had an impact on teachers' and students' practices. There was also no statistically significant difference in washback effects between male and female teachers and students, with the exception of the planning dimension, which favored female teachers, nor a statistically significant difference between less and more experienced teachers. The results also revealed a statistically significant difference in academic stream in favor of literary students.

Ghadi and Al Jamal's (2007) investigation focused on the perceptions of second-secondary English language teachers in the AL-KaraK district concerning the impact of the GSCEE on their selection of teaching techniques. A questionnaire was distributed to 92 teachers, both male, and female. The first section of the study investigated the GSCEE impact on English language teachers' practices in four areas: activity/time management, instructional methods, classroom materials, and topics teachers would teach. The study also investigated the additional GSCEE-related characteristics impact on teachers' method selection across four domains: students' learning attitudes, teachers' teaching experience, teachers' perceptions of outside pressure in the classroom, and teachers' perceptions of the significance of the GSCEE. According to the results, both positive and negative washback were found. Positively, the GSCEE had a significant impact on classroom instruction, changing how English language teachers decided to facilitate learning by making sure to provide students with feedback and involve them in activities that would prepare them for

the test. Negatively, the findings showed that teachers in AL-Karak taught English using a grammar-translation approach. These findings are in line with (Cholis & Riziq, 2018; Mniruzzman & Hoque, 2010).

In sum, the previous research on washback effects has targeted the effects on content (Hawky, 2004; Ying, 2005), teachers' methodology and classroom practices (e.g., Hawky, 2004; Smith, 1993; Wall & Alderson, 1993; Ying, 2005), syllabus and curriculum (e.g., Mniruzzman & Hoque, 2010), planning and instruction (Chen, 2002; Manjarres, 2005), and teaching material and strategies (e.g., Wall & Alderson, 1993). In addition, there are findings that confirm the coexistence of both positive and negative washback (e.g., Alderson & Wall, 1993; Ghadi & Al jamal, 2007). On the other hand, some previous studies findings indicated that changes in teachers' teaching approaches were not linked to newly introduced modifications in the test, but rather to teachers' attitudes toward these changes (e.g., Alderson & Wall, 1993; Cholis & Riziq, 2018).

3. Method

3.1 Design

This study is survey-based, descriptive in nature since surveys are one "method of studying phenomena and correctly describing them as they occur in real life and numerically expressing them" (Abbas et al., 2012, p.74).

3.2 Participants

Using convenient sampling, the researchers recruited 171 EFL teachers (86 males and 85 females) from public (n=135) and private (n=36) schools. The sample (Table 1) was comprehensive to the EFL teachers in the target directorate of education selected purposefully from one of the major cities in Jordan including 207 EFL teachers. Among this total, 30 were excluded for the purpose of the pilot study and the questionnaires of six were excluded from analysis because their responses were incomplete. Almost half of the sample were male teachers (n=86) the majority (n=135) of whom worked at public schools with an experience of more than 8 years (n=100)

3.3 Research Instrument

A 44-item, five-point Likert scale questionnaire was developed for the purpose of the current study based on a review of the extant literature (e.g., AL-Lawati, 2002; Brown, 2000; Robb & Ercanbrack, 1999; Wall & Alderson, 1993). The survey had two sections; the first elicited background information (gender, school type, and teaching experience) and the second addressed the teaching practices under three dimensions: planning (13 items), implementation (24 items), and assessment (7 items). Teachers were requested to rate the items on a five-point Likert scale (1 = never, 2 = rarely, 3 = sometimes, 4 = usually, 5 = always). That is, the higher the mean response, the greater the perceived washback effect. The questionnaire items were checked and validated by a panel of seven experts in EFL instruction: one EFL teacher, one EFL supervisor, and five faculty members specialized in TEFL.

3.4 Validity and reliability of the questionnaire

An initial 47-item questionnaire was referred to a panel of seven experts in EFL instruction to ensure comprehensiveness of the domains to the topic addressed and the comprehensive of the items to the corresponding domain in addition to item relevance to the corresponding domain, and the linguistic correctness and clarity of each single item. To ensure reliability, test-retest was applied using Cronbach's Alpha, yielding coefficient values that ranged between 0.70 to 0.83 for the dimensions with a total value of .85 for the overall scale.

3.5 Data Collection and Analysis

The participants received the questionnaires in their schools. The first researcher met with each participant individually, explained the purpose of the study, obtained consent, and distributed and collected questionnaires for analysis. Six of the collected surveys were incomplete, hence excluded from the analysis. There were 171 teacher surveys in total. The questionnaire data was analyzed using the Statistical Package for Social Sciences (SPSS) version 23 and analyzed using descriptive (mean and standard deviation values), and inferential statistics (Analysis of Variance (ANOVA and Multiple Analysis of Variance).

4. Results

4.1 JGSCEE Washback Effects on Teacher Practices on the Three Dimensions

The mean and standard deviation values for teachers' responses on the three dimensions that resulted in the following washback impacts: planning ($M = 3.83$, $SD = 0.65$), implementation ($M = 3.75$, $SD = 0.58$), and assessment ($M = 3.57$, $SD = 0.70$). Results were also considered at the level of each dimension: planning, implementation, and assessment:

As for planning, the results (Table 1) show that the mean response at the item level ranged between 3.19 and 4.13. The item about skipping listening and speaking skills elicited the highest mean response ($M = 4.13$, $SD = 1.07$). On the other hand, the item with the lowest mean response addressed allocating time for each skill based on the teachers' knowledge of JGSCEE ($M = 3.19$, $SD = 1.44$).

Table (1): Descriptive statistics for teachers' responses on planning

Item	Mean	SD	Rank	Degree
The first dimension (planning): When I plan, I...				
skip the listening and speaking skills since they are not included on JGSCEE.	4.13	1.07	1	High
modify my teaching strategies to go on with ultimate goals that meet with JGSCEE needs.	4.09	1.12	2	High
make my teaching plan associated directly with JGSCEE.	4.05	1.09	3	High
concentrate on the exercises which include grammar and vocabulary that are to be tested.	3.95	1.12	4	High
practice exam test- taking strategies.	3.89	1.14	5	High
assign much time to exercises that are comparable to JGSCEE questions.	3.85	1.17	6	High
assign time for reading and writing skills that will be tested on JGSCEE.	3.83	1.12	7	High
give grammar the first priority in my preparation since it occupies the highest share of grades.	3.81	1.32	8	High
focus on JGSCEE requirements.	3.81	1.31	8	High
dedicate much time for covering previous JGSCEE .	3.76	1.10	10	High
provide equal efforts for the four skills(listening, speaking, reading, and writing) when I prepare for JGSCEE.	3.74	1.19	11	High
depend on the student's book rather than JGSCEE questions.	3.61	1.19	12	High
allocate time for each skill (listening, speaking, reading, and writing) according to my background about JGSCEE.	3.19	1.44	13	Medium
Overall	3.83	0.65	-	High

Second, the results pertinent to implementation (Table 2) showed that the mean response at the item level ranged between 3.14 and 4.37 for the second dimension, implementation. The item concerning neglecting speaking and listening skills receives the highest mean response ($M = 4.37$, $SD = 0.86$). On the other hand, the item receiving the lowest mean response addressed covering all skills in the textbook ($M = 3.14$, $SD = 1.53$).

Table (2): Descriptive statistics for teacher responses on implementation

Item	Mean	SD	Rank	Degree
The second dimension: (implementation): When I teach, I ...				
neglect listening and speaking since they are not tested on JGSCEE.	4.37	0.86	1	Very high
apply activities which promote test- taking skills.	4.22	0.93	2	Very high
use specific teaching activities to develop my ' language skills.	4.17	0.95	3	High
greatly use previous JGSCEE papers during my teaching activities.	4.09	1.02	4	High
teach what I think is important whether it is included in JGSCEE or not.	3.99	1.17	5	High
assign my teaching activities based on their weight towards JGSCEE.	3.99	1.08	5	High
concentrate on the frequently repeated vocabulary that appear on JGSCEE.	3.96	1.11	7	High
devote much time to exam items (grammar, vocabulary) and text given higher marks on JGSCEE.	3.93	1.06	8	High
use the worksheets to review expected topics on JGSCEE.	3.89	1.18	9	High

Item	Mean	SD	Rank	Degree
assign much time to teach grammar compared to other contents (listening, speaking, reading, and writing).	3.88	1.12	10	High
concentrate on mastering the textbook's exercises that are frequently repeated on JGSCEE.	3.80	1.10	11	High
feel I am obliged to use Arabic in understanding English grammar for JGSCEE purposes.	3.76	1.07	12	High
adjust my teaching activities to meet the questions included on JGSCEE.	3.67	1.20	13	High
pay more attention to previous JGSCEE papers and stop using the textbook when JGSCEE dates are close.	3.60	1.20	14	High
skip over listening and speaking skills in the textbook because they are not tested on JGSCEE.	3.58	1.27	15	High
give equal attention to the four skills (reading, writing, listening and speaking) regardless of their weight to prepare for JGSCEE.	3.56	1.26	16	High
follow the instructions in teacher's guide in explaining the lessons whether these lessons are tested on JGSCEE or not.	3.56	1.38	16	High
teach test-taking strategies for JGSCEE purposes.	3.51	1.08	18	High
teach every section in the textbook whether it is to be tested on JGSCEE or not.	3.49	1.29	19	High
use additional commercial books because they help me to succeed on JGSCEE.	3.46	1.41	20	High
prioritize the mastery of the four skills (listening, speaking, reading, and writing) to practice the language rather than practicing JGSCEE's skills.	3.40	1.24	21	High
use software programs to help me in preparing for JGSCEE's skills.	3.25	1.24	22	High
cover all skills in the textbook (listening, speaking, reading and writing).	3.14	1.53	23	High
Overall	3.75	0.58	-	High

Third, as for assessment, the mean response (Table 3) ranged between 4.01 and 3.12. The item addressing teachers' incorporation of JGSCEE questions in their classroom examinations received the highest mean response ($M = 4.02, SD = 1.018$). On the other hand, the item with the lowest mean response ($M = 3.12, SD = 1.31$) addresses the inclusion of listening assessments in classroom quizzes.

Table (3): Descriptive statistics for teachers' responses on assessment

Item	Mean	SD	Rank	Degree
The third dimension: assessment : When I assess I				
adopt test items from JGSCEE in classroom quizzes.	4.01	1.05	1	High
assess students' assignments by using the same guidelines that are used on JGSCEE.	3.88	1.16	2	High
assess students' assignments by using the same guidelines used by teachers' grading of JGSCEE.	3.85	1.04	3	High
focus mainly on students' written and reading works.	3.63	1.29	4	High
all the task types of the textbook are covered in my made- tests.	3.53	1.31	5	High
Do not include Speaking tests in classroom quizzes and tests.	3.32	1.39	6	Medium
prepares monthly tests that mirror the content of JGSCEE rather than the content of the textbook.	3.23	1.51	7	Medium
include listening tests in classroom quizzes and tests.	3.12	1.31	8	Medium
Overall	3.57	0.70	-	High

4.2 Washback effects associated with EFL teachers' gender, teaching experience, and school type

In addition to investigating the impact of JGSCEE on EFL teachers' practices, the current study aimed at exploring the impact of some categorical variables (gender, teaching experience and school type) on teachers' reported practices.

ANOVA results (Table 4) indicate that there were no statistically significant difference ($\alpha = .05$) in teachers' responses associated with gender ($F = 0.986, P = 0.322$) or teaching experience ($F = 0.807, P = 0.448$). However, there were statistically significant differences ($F = 23.761, P = .000$) in teachers' responses based on school type.

Table (4): 3- Way ANOVA results for gender, teaching experience and school type

Source	Sum of Squares	Df	Mean Square	F	Sig.
Gender	.244	1	.244	.986	.322
Experience	.400	2	.200	.807	.448
School type	5.882	1	5.882	23.761	.000*
Error	41.094	166	.248		
Corrected Total	48.299	170			

While the above results pertain to the three dimensions of the teaching process, Multivariate Analysis of Variance (MANOVA) was used to determine whether the classifying variables (gender, teaching experience, and school type) had a significant influence on teachers' responses related to individual dimensions of the teaching process.

Concerning the first variable, gender, the findings (Table 5) show that there is no statistically significant difference ($\alpha = 0.05$) in teachers' mean response to implementation and assessment. However, there are statistically significant differences in planning favoring female teachers ($M = 3.90, SD = 0.47$) over male teachers ($M = 3.76, SD = 0.78$). In terms of teaching experience, the results show no significant differences ($\alpha = 0.05$) across all three dimensions. In terms of school type, the results show statistically significant ($\alpha = 0.05$) differences in teachers' mean response at the level of the three teaching dimensions in favor of private schools ($M = 4.17, SD = 0.63$) compared to public schools ($M = 3.73, SD = 0.62$).

Table (5): (MANOVA results) for teachers' responses by gender, teaching experience, and school type

Source	Dependent Variable	Sum of Squares	Df	Mean Square	F	Sig.
gender Hotelling's Trace=.046 F=2.498 Sig=0.062	Planning	1.539	1	1.539	4.061	.046*
	Implementation	.136	1	.136	.444	.506
	Assessment	.120	1	.120	.259	.611
Teaching experience Wilks' Lambda=.969 F=.863 Sig=0.523	Planning	1.678	2	.839	2.214	.113
	Implementation	.062	2	.031	.101	.904
	Assessment	.575	2	.288	.622	.538
School type Hotelling's Trace=.156 F=8.513 Sig=0.000	Planning	5.273	1	5.273	13.912	.000*
	Implementation	6.811	1	6.811	22.239	.000*
	Assessment	4.571	1	4.571	9.893	.002*
Error	Planning	62.919	166	.379		
	Implementation	50.842	166	.306		
	Assessment	76.708	166	.462		
Corrected Total	Planning	72.316	170			

Implementation	58.317	170
Assessment	83.709	170

5. Discussion

As the results indicate, there is a high impact for JGSCEE on EFL reported teaching practices. This is not surprising given the exaggerated significance of the JGSCEE; it is based on the exam scores students' university admission and enrollment are determined. Hence, teachers experience intense pressure from school principals, supervisors, parents, and the wider community. Pressure comes also from the media and the district administration to ensure their students' achievement of high-test scores. This pressure is evident in teachers' reported washback impact on all three dimensions of the teaching process: planning, implementation, and assessment. This finding concords with Ghadi and Al-Jamal's (2007).

At the level of the three dimensions individually, the results pertinent to planning show that teachers make plans to employ unique techniques in order to assist their students in receiving excellent grades and assist them to meet the JGSCEE testing objectives. These plans, the results indicate, focus on reading and writing at the expense of developing students' communicative competence, which means planning is test-driven. And since grammar occupies a significant portion from the JGSCEE test score, it earns teachers' attention compared to other language components. These findings are in line with those of other scholars (e.g., Alderson & Hamp-Layons, 1996; Chen, 2002; Cholis & Riziq, 2018; Manjarres, 2005). The study also shows that teachers continue to use the traditional approach to teach English, which puts more emphasis on developing students' translation and grammatical skills compared to their communicative abilities manifest in listening and speaking as reported by Caine (2005) and Wall and Alderson (1996).

Teachers seem to rarely include listening or speaking skills in their lesson plans as they are more concerned with fostering students' reading and writing skills—skills that count to scores on the JGSCEE. These findings concur with other researchers' (e.g., Caine, 2005; Haddadin, 2006; Hawkey, 2004; Mahmoudi, 2015; Manjarres, 2005; Ying, 2005) stating that newly introduced high-stakes exams force teachers to employ unique teaching strategies in order to get their interest in certain exam problems. The results of the present study, however, differ from those of Cheng (1997) who recognized that changes in terms of the test format did not change teachers' plans or teaching methods. Changing the exam question-format, therefore, might not motivate teachers to change their teaching practice.

With regard to implementation, the present study indicates that teachers' practices in the classroom are highly affected by JGSCEE since they tend to neglect listening and speaking skills, teach their students test-taking skills, and provide them with previous exam papers. In addition, EFL teachers give special attention to teaching grammar and vocabulary. This finding is confirmed by the result that—according to teachers' responses—the item with the lowest mean value (medium) was related to allocating time for each language skill. The conclusion is that teachers assign weight for test-oriented skills, namely reading and writing. This finding is similar to others' findings (e.g., Koshima et al, 2018; Maniruzzaman & Hoque, 2010; Mnjarres, 2005; Ying, 2015).

With pertinence to assessment, teachers' practices are also highly affected by JGSCEE guidelines. The results of this study show that teachers adopt test items from previous JGSCEE in their quizzes and assess students' assignments according to the guidelines adopted in grading JGSCEE papers, a practice that might be interpreted by teachers' attempt to familiarize students with JGSCEE exam question format and grading policy. Also, listening and speaking tests are neglected in teachers' exams. A plausible explanation for this is that JGSCEE does not include listening and speaking tasks; thus, teachers tend to overlook these skills in their planning, implementation and assessment. Moreover, it should be noted that EFL teachers' performance in the eyes of school principals is not independent of students' grades. Towards this end, teachers are tempted to teach for the test and solely for the test. This finding is consistent with the findings of other researchers (e.g., Manjarres, 2005; Shohamy et al., 1996; Tayeb et al., 2014; Wall & Alderson, 1993). Overall, these findings lend support to Alderson and Wall's theory that stated that high stake-exam affect what and how teachers teach. They also go in line with those of earlier investigations (e.g., Alderson & Hamp-Lyons, 1997; Al-lwati, 2002).

As for the impact of classifying variables, the findings of the study reveal a statistically significant difference in washback effects of the JGSCEE on teachers' practices that are associated to gender in favor of female teachers in association with planning. This might be explained on the ground that female teachers are more motivated to plan lessons than male teachers because they want to avoid criticism from their principals and supervisors, and they are more committed to their paperwork. As reported by Al-lawati (2002), female teachers are more influenced by the JGSCEE in planning their lessons. El-Emadi et al. (2019) and Taqi et al. (2015) reported more attention to planning by female teachers.

Another finding that is related to the teaching experience; the study findings suggest that novice teachers and experienced teachers behave alike. This might be attributed to the high impact of this exam, which makes both experienced and less

experienced teachers equally concerned with the results of their students. It seems unclear where experience stands in terms of teachers' practices, for whereas some (e.g., Al-lawati, 2002) reported equal impact, others (e.g., Shohamy et al., 1996) reported that experienced teachers were influenced more than novice teachers in their teaching methods and assessment. This lack of agreement on the impact of experience requires further research.

Our findings suggest differences associated with the school type--with a higher impact on private school teachers at the level of the three instructional dimensions: planning implementation, and assessment. One possible justification for this finding amounts to accountability; private schools' reputation and funding is tied strongly to students' results, and teachers might easily get a contract termination based on students' scores, a practice that probably never took place in the public sector.

6. Conclusion

The current study concludes that EFL teachers experience a high washback effect of the JGSCEE on their practice as they seem undergo intense pressure from their principals, supervisors, parents, and the larger community in addition to the public media to ensure that their students achieve high scores. This effect is evident on the three dimensions of the teaching process: planning (e.g., Chen, 2002; Cholis & Riziq, 2018; Manjarres, 2005), implementation (e.g., Maniruzzaman & Hoque, 2010; Manjarres, 2005; Ying, 2015), and assessment (e.g., Manjarres, 2005; Wall & Alderson, 1993). EFL teachers' practices as they go under pressure seem to be not totally independent of their gender with a higher commitment in favor of female teachers, which seems to be culturally bound. The pressure high-stake exams put on teachers' shoulders seems to lead them to run behind what brings higher scores at the expense of building and enhancing students' linguistic capacity, marginalizing the impact of experience, and accountability is likely to play an important role in shaping teachers' practices but not necessarily guarantee deep understanding that goes beyond obtaining higher scores.

Educational policymakers are invited to reconsider the implementation of the JGSCEE as a sole representative of the entire assessment realm and a student's score as the only ground for determining a student's admission to university. This is likely to eliminate the test-driven doctrine in the teaching process. It is then that EFL teachers would designing the test format in such a way that it covers the four skills integrated and assesses students' higher-level thinking and problem-solving abilities, as well as their communicative ability. EFL teachers should strike a balance between the need to prepare their students for high-stake exit exams and the need to improve their students' language competence. EFL teachers should also follow the guidelines in the teacher's book to ensure the development of all language skills. Furthermore, teachers should concentrate on improving students' speaking and listening skills. Thus, it is legitimate to ask: Through high-stake exams, are we producing language teachers or test teachers?

Finally, it should be noted that this study is not devoid of limitations. The number of the participants is a limitation, and the data collection instrument is another. Future research may include a higher number of participants and involve more instruments towards more in-depth investigation.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers.

References

- [1] Abbas, M. K, Nofal, M. B., Mustafa, M., & Awad, F.H.A (2012). *An introduction to research methods in education and psychology*. Amman, House of the March.
- [2] Al-Lawati, N. (2002). *Washback effect of secondary certificate English examination on teaching and learning processes* (Doctoral dissertation, Sultan Qaboos University).
- [3] Bachman, F., & Palmer, A. (2000). *language testing in practice*. Oxford: Oxford University Press.
- [4] Baker, E. L., & O'Neil Jr, H. F. (1994). Performance assessment and equity: A view from the USA. *Assessment in Education: Principles, Policy & Practice*, 1994, 1.1: 11-26.
- [5] Barners, M. M. (2016). The washback effects of the TOEFL iBT in Vietnam. *Australian Journal of Teacher Education*, 41(7), 158-174.
- [6] Alderson, J. C., & Wall, D. (1993). *Does washback exist? Applied linguistics*, 14(2), 115-129.
- [7] Chen, L. M. D. (2002). *Taiwanese junior high school English teachers' perceptions of the washback effect of the basic competence test in English* (Doctoral dissertation, The Ohio State University).
- [8] Cheng, L. (2005). *Change language teaching through language testing*: Cambridge University Press.
- [9] Cholis, H. W. N., & Riziq, F. (2018). Senior high school English teachers' Perceptions on a High- Stakes Test (SBMPTN): A Washback Study. *International Journal of Education and Literacy Studies*, 6(3), 47-52.
- [10] Eckstein, M. A., & Noah, H. J. (1993). *Secondary school examinations: International perspectives on policies and practice* (pp.243-245). New Haven, CT: Yale University Press.
- [11] El-Emadi, A.A., Said, Z., & Friesen, H. L. (2019). Teaching style differences between male and female science teachers in qatari schools: Possible impact on student achievement. *EURASIA Journal of Mathematics, Science and Technology Education*, 15(12), 35-47

- [12] Frederickson, J. R. & Collins, A. (1998) . A system approach to educational educational testing. *Educational researcher*, 18(9), 27-32.
- [13] Ghadi, N., & Al-Jamal, D. (2008). English language general secondary certificate examination washback in Jordan. *Asian EFL Journal*, 10(3), 158-186.
- [14] Hawkey, R. O. G. E. R. (2004). An IELTS Impact Study: Implementation and some early findings. *Research Notes*, 15 (3), 12-16.
- [15] Haddadin, A. (2006). *A remedial program for developing the English language skills in the light of the washback effect of public examinations on the instruction of the secondary stage in Jordan*. (635169)[Doctoral dissertation, Amman Arab University for Graduate studies].DAR ALMANDUMAH.
- [16] Kirkpatrick, R., & Gyem, K. (2012). Washback effects of the new English assessment on secondary schools in Bhutan. *Language testing in Asia*, 2(4), 1-28.
- [17] Maniruzzaman, M., & Hoque, M. E. (2010). How does washback work on the EFL syllabus and curriculum? A case study at the HSC level in Bangladesh. *Language in India*, 10(12), 14- 25.
- [18] Manjarres, N. B. (2005). Washback of the foreign language test of the state examinations in Colombia: A case study. *Journal of Second language Acquisition and Teaching*, 12 (3), 1-19.
- [19] Messick, S. (1996).Validity and washback in language testing. *Language testing*, 13(3), 241-256.
- [20] Ministry of Education. (2014, October). Educational Statistical Report. Amman: MOE. Retrived from <http://www.moe.gov.jo/MenuDetails.aspx?MenuID=29>.
- [21] Ozmen, K. (2011). Washback effects of the inter-University foreign language examination on foreign language competences of candidate academics. *Novitas Royal (Research on Youth and language)*, 5(2), 128-137.
- [22] Prodromou, L. (1995). The Backwash Effect: From Testing to Teaching. *ELT Journal*. 49(1), 13-25.
- [23] Popham, W. J. (1987). The merits of measurement – driven instruction .The Phi Delta Kappan, 68(9), 679-682.<http://www.jstor.org/stable/20403467>
- [24] Ruder, L., & Schafer, W. (2002). *What Teachers Need to Know about Assessment*. Washington: National Education Association.
- [25] Shatanawi, M. (2005). Describing an EFL Integrative Programme based on the Whole Language Approach and Measuring its effects on the Basic Stage ' Achievements in English . Unpublished doctoral dissertation , Amman Arab University for Graduate Studies , Jordan.
- [26] Shohamy, E. (2020). *The power of tests : A critical perspective on the uses of language tests* . Routledge.
- [27] Shepard, L. A. (1993). The place of testing reform in educational reform: A reply to Cizek. *Educational Researcher*, 22(4), 10-13.
- [28] Tsagari, D. (2007). *Investigating the Washback effects of the FCE examination* (Cambridge/ESOL) in Greece. Greece, UCLES.
- [29] Wall, D. (2005). The Impact of High-stakes Examinations on Classroom Teaching: A Case Study Using Insights Testing and Innovations Theory, *Studies in Language Testing* 22, UK: Cambridge University Press.
- [30] Wang, J. (2010). A study of the Role of the Teacher in Washback. Unpublished PhD thesis, McGill University, Canda.
- [31] Ying, Y. (2005). Washback Effects of the Spoken English Test on Textbook Content . *University of China publication*, 7(4), 21-32.