

Teachers' Perceptions of Washback Effects of Standardized Examinations in Intermediate and Secondary Schools in Kuwait

Amani Manea Alajmi, Rasha Alaudan
King Saud University, Riyadh, Saudi Arabia

This study investigated teachers' perceptions of the washback effects of English standardized examinations in intermediate and secondary public schools in Kuwait. The study aimed to examine how standardized examinations influence teaching content, teaching methods, and classroom assessment practices. A quantitative research design was adopted, and data were collected through a questionnaire administered to 92 English language teachers. The findings indicated that standardized examinations in Kuwait generate both positive and negative washback effects, highlighting the complexity of washback as a phenomenon. Positive washback was reflected in teachers' perceptions that standardized examinations support the development of reading and writing skills, help teachers align instruction with curriculum objectives, identify students' learning gaps, adapt teaching methods, and maintain attention to all four language skills during classroom assessment. At the same time, the findings revealed several forms of negative washback. Teachers reported greater emphasis on exam-related content and task types, test-taking techniques, and examination readiness. Negative washback was also reflected in administrative pressure to improve examination scores, students' focus on grades rather than language development, and the tendency to copy examination formats in classroom assessment.

Keywords: washback, standardized examinations, high-stakes testing, ESL, EFL, language assessment

Introduction

Standardized examinations and their washback effects have become significant issues in English as a Second Language (ESL) and English as a Foreign Language (EFL) education because such examinations can influence both teaching and learning practices. One common form of assessment used in many educational systems is standardized testing, which refers to tests that are administered and scored using systematic and consistent procedures (Brown & Abeywickrama, 2019). Standardized examinations are widely used in schools around the world to assess students' mastery of the standards or competencies that are prescribed for specific grade levels. They are also used to determine whether students meet requirements for graduation or for entry into higher levels of education (Brown & Abeywickrama, 2019).

Washback, which Alderson and Wall (1993) defined as the impact that testing has on teaching and learning, is commonly associated with standardized and high-stakes examinations in language education. Some researchers argue that examinations can be used to bring about positive changes in teaching practices, which is referred to as

Amani Manea Alajmi, Master's program of TESOL, Department of Curriculum and Instruction, College of Education, King Saud University, Riyadh, Saudi Arabia.

Rasha Alaudan, Ph.D., Assistant Professor, Department of Curriculum and Instruction, College of Education, King Saud University, Riyadh, Saudi Arabia.

positive washback (Cheng & Curtis, 2004). On the other hand, negative washback occurs when tests do not reflect course objectives or learning principles (Cheng & Curtis, 2004). The growing dependence on standardized and high-stakes examinations has made testing a powerful factor in language education (Al-Wadi, 2020). In response to the requirements of these examinations, EFL teachers have been required to modify their instructional approaches so that students can meet the competencies expected in such tests (Al-Wadi, 2020). Furthermore, many EFL teachers report that the implementation of standardized English language examinations has led them to adopt particular teaching methods when teaching the target language to ESL/EFL learners (Al-Wadi, 2020). In addition, the increasing demand for standardized English language tests can lead to a shift away from the primary goal of learning English (Choi, 2008). Instead, the focus may become mainly achieving the highest possible score on EFL examinations (Choi, 2008).

Wall and Alderson (1996) identified three main areas in which washback can manifest: teaching content, teaching methods, and assessment practices. In terms of content, teachers may rely on materials that do not follow the official curriculum or concentrate only on topics likely to appear in the examination (Wall & Alderson, 1996). With regard to teaching methods, instruction may emphasize exam-related techniques rather than the methods recommended in the curriculum (Wall & Alderson, 1996). Finally, washback may also affect classroom assessment, where tests and quizzes closely resemble past examination papers (Wall & Alderson, 1996).

Recent research highlights the development of washback studies. Xie and Jia (2025) conducted a bibliometric analysis of 243 empirical studies published between 1993 and 2024. Their findings indicate that washback research has expanded significantly over time and gained increasing scholarly attention (Xie & Jia, 2025). Their study also shows that research on washback is particularly prominent in Asian contexts, where centralized education systems often rely on high-stakes standardized examinations (Xie & Jia, 2025). This emphasizes the importance of investigating washback in Asian EFL contexts such as Kuwait, where high-stakes standardized testing is used.

Despite the growing body of research, there is limited research investigating how standardized examinations influence English language teaching and learning in Kuwait, particularly across both intermediate and secondary levels in public schools. Moreover, there is a lack of research that systematically examines the specific areas of washback identified by Wall and Alderson (1996). Therefore, this research aims to address this gap by investigating the washback effect of standardized examinations in Kuwait across intermediate and secondary public education, with a particular focus on teaching content, teaching methods, and classroom assessment practices.

In addition, this study investigates teachers' perceptions of washback. Teachers' beliefs and perceptions can play a significant role in shaping how examinations influence classroom practices. Empirical evidence supports this relationship. For instance, Liang, Siraj, and Lo (2025) conducted qualitative interviews with four experienced high school English teachers and found that washback on teaching practices was positive when teachers' beliefs were aligned with the assessment. However, when teachers' beliefs were not consistent with the examination, the washback effects were either negative or insignificant.

Therefore, the objectives of this study are to:

1. Investigate teachers' perceptions of washback effects resulting from standardized examinations in intermediate and secondary schools in Kuwait.
2. Examine how standardized examinations influence teaching content in intermediate and secondary classrooms.

3. Explore how standardized examinations affect teaching methods.
4. Investigate the influence of standardized examinations on classroom assessment practices.

This study seeks to address the following research questions:

1. How do teachers perceive the washback effects of standardized examinations in intermediate and secondary schools in Kuwait?
2. How do standardized examinations influence teaching content in intermediate and secondary classrooms?
3. How do standardized examinations affect teaching methods?
4. How do standardized examinations influence classroom assessment practices?

Literature Review

Context of the Study

English language education in Kuwait is a central component of the government school system. English is taught continuously from Grade 1 to Grade 12, which provides students with twelve years of formal English instruction in public schools (Tryzna & Al Sharoufi, 2017). At the national level, English language education is supervised by the English Language Teaching (ELT) General Supervision Department within the Ministry of Education. This department is responsible for overseeing English language instruction and assessment across the country (Tryzna & Al Sharoufi, 2017). Its responsibilities include establishing national standards for English language education, developing the curriculum, preparing pedagogical materials such as textbooks and workbooks, recommending teaching methods, and producing and administering assessment tools (Tryzna & Al Sharoufi, 2017). In addition, the ELT General Supervision department determines the learning objectives for each educational level and evaluates learning outcomes (Tryzna & Al Sharoufi, 2017). It also supervises the preparation of English teachers for primary, intermediate, and secondary schools (Tryzna & Al Sharoufi, 2017).

The English language curriculum in Kuwait views English education not only as the learning of a school subject but also as a means of developing students' ability to communicate effectively, think critically, and participate responsibly in an interconnected world (Ministry of Education, 2026). Within this framework, English is considered an important tool for communication in a globalized and technologically advanced world. Proficiency in English is therefore associated with opportunities for academic, personal, and professional development (Ministry of Education, 2026).

Accordingly, English language education is intended to move beyond rote memorization and textbook-centered instruction toward learning experiences that address the communicative needs of learners in real-life contexts (Ministry of Education, 2026). The curriculum integrates the development of the four language skills—listening, speaking, reading, and writing—together with broader competencies such as expression, intercultural awareness, and preparation for participation in fields such as science, technology, business, and education (Ministry of Education, 2026). Therefore, the curriculum draws on several established approaches in language teaching. These include Communicative Language Teaching, which emphasizes authentic communication and negotiation of meaning; and Task-Based Language Teaching, which promotes learning through meaningful tasks such as interviews and projects (Ministry of Education, 2026).

At the secondary level, the English curriculum emphasizes the integration of language skills through tasks and activities for both learners and teachers. The learning outcomes presented at the beginning of each module are intended to support this integration and promote interaction between teachers and students (Ministry of Education, 2026). Students at this stage are expected to demonstrate greater fluency and confidence in spoken

English, enabling them to engage in more demanding and realistic communicative situations with unpredictable or open-ended outcomes. As a result, the course includes more group work and discussion activities than in earlier levels (Ministry of Education, 2026). Students are also expected to demonstrate broader cognitive understanding and greater independence in learning, as they are approaching adulthood. For this reason, the curriculum encourages learners to express their ideas, take greater responsibility for their learning, and collaborate more actively with their peers (Ministry of Education, 2026).

Assessment also plays a significant role in the Kuwaiti education system. In the intermediate stage, final standardized examinations constitute 60% of students' total grades, and students sit for two examinations each academic year, one at the end of the first semester and another at the end of the second semester (Ministry of Education, 2026). In the secondary stage, final standardized examinations account for 70% of the total grades, with students similarly taking two examinations per year, one in each semester (Ministry of Education, 2026). These examinations therefore represent a major component of student evaluation within the national education system.

In Kuwaiti public schools, English examinations in both intermediate and secondary stages assess multiple language components. At the intermediate level, the exam includes reading, vocabulary, grammar, and writing, with reading tasks based on an unseen text that tests comprehension skills such as identifying the main idea, understanding vocabulary, locating explicit and implicit information, and recognizing the writer's purpose (Ministry of Education, 2026). The writing section evaluates grammar knowledge and requires students to produce a short paragraph with a clear topic sentence, supporting details, and a conclusion (Ministry of Education, 2026). At the secondary level, the summative examinations assess a wider range of components, including vocabulary, grammar, writing, reading comprehension, summary writing, and translation, as well as language functions that involve responding to communicative situations and set-book questions that consist of productive questions of a general nature based on the prescribed textbook content (Ministry of Education, 2026).

Standardized Examination

Standardized examinations refer to tests that are developed through systematic procedures to ensure consistency in their design, administration, and scoring (Brown & Abeywickrama, 2019). Phelps (2005) described standardized testing as a systematic method of collecting information that supports educational decision-making. In Kuwait, the English examination system can be considered standardized. Tryzna and Al Sharoufi (2017) explained that, in Kuwait, students' progress and the achievement of learning objectives are assessed annually through national English language examinations that are prepared, administered, and evaluated by the Ministry of Education through ELT General Supervision. This centralized preparation and administration reflect the key characteristics of standardized testing. Although these examinations are commonly described as summative assessments in Kuwait, this classification does not contradict their standardized nature. Mansory (2016) explained that summative assessments are often standardized, meaning that students at the same level take identical tests and are evaluated using consistent scoring procedures.

Standardized examinations are characterized by fairness and consistency in assessing student performance. Sireci (2005) explained that a standardized test is a test administered under uniform conditions, with equivalent content and consistent scoring procedures for all test takers. Standardization is intended to promote fairness by ensuring that everyone takes the same test under the same conditions (Sireci, 2005). Sireci (2005) argued that standardization itself does not make a test unfair, biased, or focused only on unimportant knowledge.

Standardized tests are designed to provide objective information (Sireci, 2005). He also rejected two common misconceptions: that standardized tests contain only multiple-choice questions and that they measure only lower-level thinking skills (Sireci, 2005).

Phelps (2005) is considered one of the scholars who highlighted the benefits of standardized examinations. He argued that standardized tests can provide diagnostic information by identifying students' strengths and weaknesses (Phelps, 2005). Phelps (2005) also emphasized that test results can be used to predict future performance such as college admission, scholarships, and employment opportunities. However, McNeil (2000) criticized the impact of standardized examinations on education. She argued that a strong emphasis on standardized assessments may narrow the curriculum and reduce the depth and quality of teaching (McNeil, 2000). Teachers may increasingly focus on test-related content and preparation activities, which can replace broader and more meaningful learning experiences (McNeil, 2000). She also noted that such systems may reduce teachers' authority over classroom practices (McNeil, 2000).

Washback

Washback, also known as backwash, has been widely defined as the influence that testing exerts on teaching and learning (Alderson & Wall, 1993; Hughes, 2003; Bachman & Palmer, 1996). The work of Alderson and Wall (1993) is frequently regarded as a foundational contribution to washback research. They defined washback as the degree to which a test influences teachers and students to do things they would not normally do (Alderson & Wall, 1993). To clarify the concept of washback, they proposed 15 hypotheses that distinguish between effects on attitudes, content, methods, and processes in both teaching and learning (Alderson & Wall, 1993). These hypotheses suggest that tests may influence teaching and learning, what and how teachers teach, what and how learners learn, as well as the rate, sequence, and depth of teaching and learning (Alderson & Wall, 1993). They further proposed that tests that have important consequences are more likely to produce washback, and that such effects may vary across individuals (Alderson & Wall, 1993).

The distinction between positive and negative washback has been emphasized in subsequent literature. According to Cheng and Curtis (2004), positive washback refers to situations in which changes in examinations lead to beneficial changes in teaching and learning. In such cases, teachers and learners demonstrate positive attitudes toward the test and work collaboratively toward its objectives (Cheng & Curtis, 2004). In contrast, negative washback refers to the undesirable influence of tests on teaching and learning (Alderson & Wall, 1993). Alderson and Wall (1993) described negative washback as the effect of a test that leads teachers or learners to engage in activities they do not wish to teach or learn. In such situations, tests may not align with learning principles or course objectives, leading teaching and learning to focus mainly on the test, even if teachers and students do not fully support or understand it (Cheng & Curtis, 2004).

Alderson and Wall (1993) further emphasized that the relationship between tests and their effects is complex. They argued that the quality of washback is not necessarily dependent on the quality of the test (Alderson & Wall, 1993). Both good and bad tests may produce either positive or negative effects, indicating that washback is not determined solely by test quality (Alderson & Wall, 1993). Furthermore, Wall and Alderson (1996) identified three main areas in which washback can occur: teaching content, teaching methods, and assessment practices. Washback in teaching content may involve reliance on materials that do not follow the official curriculum or a focus on topics likely to appear in examinations (Wall & Alderson, 1996). In terms of teaching methods, instruction may prioritize exam-related techniques rather than recommended pedagogical approaches (Wall &

Alderson, 1996). In classroom assessment, washback may be reflected in tests and quizzes that closely resemble past examination papers (Wall & Alderson, 1996).

Empirical evidence from a study conducted by Wall and Alderson (1996) in Sri Lanka further illustrates these effects. This study examined the impact of the ESL examination on classroom instruction using classroom observations, interviews, questionnaires, and test analyses (Wall & Alderson, 1996). The findings indicated that washback occurred in both positive and negative forms (Wall & Alderson, 1996). In terms of teaching content, teachers demonstrated efforts to cover textbooks, which was generally positive, although it became negative when it involved memorization of texts likely to appear in examinations (Wall & Alderson, 1996). Additional negative washback effects included increased focus on reading, reduced attention to oral skills, and the exclusion of listening lessons due to their absence from the examination (Wall & Alderson, 1996). Extended periods of exam preparation also led to the neglect of speaking and listening activities (Wall & Alderson, 1996). Regarding teaching methods, the study found no evidence of washback (Wall & Alderson, 1996). This was attributed to the lack of guidance provided by the examination and teacher guides on how to teach specific skills (Wall & Alderson, 1996). As a result, teaching methods remained unchanged (Wall & Alderson, 1996). In contrast, washback was evident in assessment practices (Wall & Alderson, 1996). Positive washback effects included greater emphasis on reading and writing and the use of exam-related question types when aligned with textbook content (Wall & Alderson, 1996). Negative washback effects included the use of exam question types not found in textbooks and the direct copying of materials from past examination papers (Wall & Alderson, 1996).

Previous studies on washback. Cheng (1997) examined the washback effect of the revised Hong Kong Certificate of Education Examination (HKCEE) in English on teaching practices in secondary schools in Hong Kong. Using a mixed-methods design that included questionnaires, interviews, and classroom observations, the study examined how changes in the examination influenced teaching and learning (Cheng, 1997). Participants included teachers, students, and other stakeholders such as textbook publishers (Cheng, 1997). The findings showed that the revised examination produced noticeable washback on classroom activities, teaching materials, and teaching practices (Cheng, 1997). Teachers reported greater integration of language skills and more interactive activities but also increased pressure due to the need to adapt to the new exam format (Cheng, 1997). Teaching materials changed rapidly, as publishers revised textbooks and provided new resources to match the examination. However, changes in teaching methods were slower and more difficult than changes in teaching materials. The study concluded that there was insufficient evidence to determine whether the washback effects were positive or negative (Cheng, 1997).

Azadi and Gholami (2013) examined the washback effect of English language tests on teaching in Iranian high schools by analyzing examination papers, observing classrooms, and collecting questionnaire data from teachers and students. The results showed a strong relationship between what was taught in classrooms and what final tests measured, indicating a clear washback effect (Azadi & Gholami, 2013). However, this washback was found to be negative. Teaching was restricted by examination demands, which did not ensure the teaching of all required skills such as listening, speaking, and writing (Azadi & Gholami, 2013). Students were just involved in comprehension and structure tasks which are different from what is outlined in the textbooks, syllabus, and curriculum assigned by the Ministry of Education of Iran (Azadi & Gholami, 2013). Overall, the study concluded that the tests narrowed the curriculum and reduced students' opportunities to develop full linguistic and communicative competence (Azadi & Gholami, 2013).

Das, Shaheen, Shrestha, Rahman, and Khan (2014) explored the alignment between curriculum reform and assessment practices in Bangladesh through a mixed-methods study involving 38 schools. The findings revealed a mismatch between curriculum objectives and assessment practices. Most teachers were dissatisfied with the assessment system as it focused mainly on reading and writing, while listening and speaking were excluded (Das et al., 2014). The analysis of question papers also showed that test items required little student engagement and were not communicative in nature (Das et al., 2014). As a result, teachers prioritized exam preparation and memorization rather than communicative skill development (Das et al., 2014). The study concluded that the lack of alignment between curriculum and assessment produced negative washback and limited the effectiveness of educational reform (Das et al., 2014).

Aron (2015) investigated teachers' perceptions of the washback effect of the Estonian national school-leaving examination in English. The study aimed to clarify the washback effect of the examination and tested three hypotheses: that the format and content of the examination affect teaching, that the examination shapes what teachers teach, and that there is more evidence of negative washback than positive washback (Aron, 2015). The first two hypotheses were verified. The study confirmed that the format and content of the examination influenced teaching in several ways, including narrowing the curriculum to exam-related tasks, emphasizing exam formats, and using mock exams (Aron, 2015). The study also found that the examination shaped what teachers taught, as teachers began to give greater attention to exam content, particularly writing and speaking tasks, as well as oral exam simulation activities (Aron, 2015). However, the third hypothesis was not verified, as the study found it difficult to determine whether the observed effects were positive or negative (Aron, 2015). Some practices could be seen as beneficial because they supported exam success and speaking development, yet at the same time they could be viewed as harmful when caused by examination pressure (Aron, 2015).

Al Amin and Greenwood (2018) investigated the role of examinations in Bangladesh secondary education and their impact on teaching and learning. The study presented findings showing a gap between teachers' perceptions of curriculum expectations and their reported practices, as well as findings on the influence of the examination system (Al Amin & Greenwood, 2018). The study found that teachers claimed they tried to teach all textbook chapters, but as the examination approached, they focused on content considered more important for the examination (Al Amin & Greenwood, 2018). Some also reported that time limitations prevented them from covering all chapters and led them to ignore chapters less likely to appear in the examination (Al Amin & Greenwood, 2018). The findings suggested that language learning in this context became the acquisition of correct answers rather than the development of communication and thinking (Al Amin & Greenwood, 2018). Students could pass examinations without necessarily understanding meaning or being able to use English in real-life contexts (Al Amin & Greenwood, 2018). In addition, many teachers recognized the importance of listening and speaking but did not include them in their teaching because they were not tested (Al Amin & Greenwood, 2018). The study concluded that the examination system reduced academic curiosity and limited the achievement of curriculum objectives (Al Amin & Greenwood, 2018).

Al-Wadi (2020) examined the influence of standardized national English examination policy on the teaching practices and beliefs of secondary schools EFL teachers in Bahrain. The study used a mixed-methods design, with questionnaires and interviews involving 60 teachers (Al-Wadi, 2020). The findings indicated that the examination influenced teaching in several ways. It directed teaching toward specific learning outcomes and exam-related requirements, while limiting teachers' ability to implement innovative methods, design engaging lessons, and respond to students' individual needs (Al-Wadi, 2020). At the same time, some participants believed

the examination helped them teach new skills and design effective teaching materials (Al-Wadi, 2020). In addition, teachers perceived the examination as guiding their assessment practices (Al-Wadi, 2020). Overall, teachers aligned their instruction with exam requirements and felt responsible for preparing students for the exam (Al-Wadi, 2020). The study concluded that the examination functioned both as a source of development and as a constraint (Al-Wadi, 2020).

Rahman, Jamila, and Siam (2024) examined teaching practices and language skill emphasis at the Higher Secondary Certificate level in Bangladesh through surveys and interviews with students and teachers. Even though the curriculum emphasized effective communication through the four language skills, the study found that listening and speaking were neglected in teaching, curriculum implementation, and assessment (Rahman et al., 2024). Reading and writing received much greater emphasis (Rahman et al., 2024). Both teachers and students acknowledged the importance of all language skills but reported dissatisfaction with current practices (Rahman et al., 2024). The study identified a clear mismatch between assessment and curriculum objectives, concluding that this lack of alignment resulted in poor language proficiency (Rahman et al., 2024).

Huang and Zeng (2025) investigated the sustainability of washback effects following the reform of China's National Matriculation English Test, which changed from a once-yearly to a twice-yearly high-stakes examination. The study examined the immediate and long-term effects of the reform on teachers' beliefs and practices using a hybrid longitudinal and cross-sectional design (Huang & Zeng, 2025). The findings showed that the reform produced significant washback on teachers' beliefs and practices. Immediately after the reform, teachers changed their planning, teaching content, and methods to match new test requirements (Huang & Zeng, 2025). There was increased emphasis on writing, integrated tasks, exam-related activities, and frequent classroom tests (Huang & Zeng, 2025). Teachers also reported heavier workload and greater pressure because of the increased demands of preparation and marking (Huang & Zeng, 2025). Over time, some washback effects decreased, especially teachers' reliance on exam-oriented strategies and their sense of pressure, as they gradually adapted to the reform (Huang & Zeng, 2025). However, the examination continued to influence teaching planning, workload, and professional development even four years after implementation (Huang & Zeng, 2025).

Li and Hu (2025) investigated the washback effect of the Chinese Proficiency Test on teaching and learning. A mixed-methods design was used, combining questionnaires, classroom observations, and semi-structured interviews involving teachers and students (Li & Hu, 2025). The findings indicated that the test produced a relatively strong positive washback effect on both teaching and learning, although some unintended and negative effects were also observed (Li & Hu, 2025). For teachers, the examination clarified instructional objectives and affected teaching content, methods, and attitudes (Li & Hu, 2025). Teachers frequently incorporated test-taking strategies into their lessons and organized their teaching around test requirements (Li & Hu, 2025). Lecture-based teaching was used most often, while communicative activities were less common because of time limitations and exam preparation demands (Li & Hu, 2025). Despite this, teachers generally maintained positive attitudes and became more attentive and rigorous (Li & Hu, 2025). For students, the examination influenced learning content, strategies, and attitudes (Li & Hu, 2025). Students often focused on vocabulary memorization, textbook study, and practice of previous test questions (Li & Hu, 2025). However, they were not limited to mechanical learning as they gradually adopted more meaningful cognitive and metacognitive strategies (Li & Hu, 2025). Overall, the study concluded that the Chinese Proficiency Test strongly shaped both teaching and learning practices, producing both beneficial and limiting washback effects (Li & Hu, 2025).

Methodology

Research Design

This study adopts a quantitative research design using a survey method to investigate teachers' perceptions of the washback effects of standardized examinations in intermediate and secondary schools in Kuwait. A quantitative approach is suitable as it allows for the systematic and consistent collection and analysis of numerical data to identify patterns and trends in participants' responses (Cohen, Manion, & Morrison, 2018). As noted by Creswell (2009), quantitative designs are particularly appropriate for examining attitudes and perceptions across relatively large samples.

Research Context

This study is conducted within the context of public schools in Kuwait, with a focus on the intermediate and secondary levels, where standardized examinations are implemented. These examinations are developed and administered by the Ministry of Education and play a significant role in determining students' academic outcomes. This context provides a suitable setting for examining the washback effects of such examinations on teachers' practices.

Participants

The target population consists of English language teachers working in intermediate and secondary public schools in Kuwait. A snowball convenience sampling method was used in this study. This sampling method involves choosing participants based on their availability and willingness to participate and then asking them to suggest other eligible participants (Cohen et al., 2018). A total of 92 teachers participated in the study. The demographic characteristics of the participants were described according to gender, years of teaching experience, and teaching level, as shown in Table 1.

Table 1

Demographic Characteristics of the Study Sample (N = 92)

Variable	Category	N	%
Gender	Male	35	38.0%
	Female	57	62.0%
Years of teaching experience	0-5 years	26	28.3%
	6-10 years	23	25.0%
	More than 10 years	43	46.7%
Teaching level	Intermediate	65	70.7%
	Secondary	27	29.3%

Research Instrument

Data were collected using a structured questionnaire. The questionnaire consisted of 24 items distributed across four dimensions as follows:

Dimension 1, *Teachers' general perceptions of washback effects* (Items 1-6).

Dimension 2, *The influence of standardized examinations on teaching content* (Items 7-12).

Dimension 3, *The influence of standardized examinations on teaching methods* (Items 13-18).

Dimension 4, *The influence of standardized examinations on classroom assessment practices* (Items 19-24).

The questionnaire included both positively and negatively worded items. The negatively worded items were Items (5-6), (9-12), (15-18), and (21-24). Negatively worded items were reverse-coded prior to analysis so that

higher scores consistently indicated more positive perceptions of washback effects. Dimension scores were computed as mean scores for the items included in each dimension.

The questionnaire was mainly developed by the researcher based on the theoretical framework of washback proposed by Wall and Alderson (1996), along with other theoretical framework on the washback effects of standardized examinations. In addition, some items were adapted from previous research conducted by Al Amin and Greenwood (2018). These items were selected due to their relevance to the research objectives and their alignment with the context of standardized examinations. The adapted items were modified where necessary to ensure clarity and suitability for the Kuwaiti educational context.

A pilot study was conducted with a small group of teachers to check the clarity of the questionnaire items and instructions, identify any wording issues, and determine the suitability of the items (Cohen et al., 2018). The feedback indicated that the items were clear and appropriate, and no revisions were required. The responses obtained from the pilot study were excluded from the final analysis.

Questionnaire scoring. The questionnaire items were rated on a five-point Likert scale ranging from Strongly Agree to Strongly Disagree. This format allows for a degree of sensitivity and variation in response while still producing quantitative data (Cohen et al., 2018). Responses were scored as follows: Strongly Agree = 5, Agree = 4, Neutral = 3, Disagree = 2, and Strongly Disagree = 1. After reverse-coding the negatively worded items, higher scores indicated more positive perceptions of washback effects, whereas lower scores indicated less positive perceptions.

To interpret the mean scores of the questionnaire items and dimensions, the range of the scale was calculated, and the mean scores were interpreted as shown in Table 2.

Table 2

Mean Score Ranges and Their Interpretation

Mean range	Interpretation
4.20-5.00	Strongly agree
3.40-4.19	Agree
2.60-3.39	Neutral
1.80-2.59	Disagree
1.00-1.79	Strongly disagree

Validity and reliability of the questionnaire. Validity. Internal consistency validity was examined by calculating Pearson correlation coefficients between each item and the total score of the dimension to which it belongs. Table 3 shows that all item-dimension correlations were positive and statistically significant, indicating satisfactory internal consistency for all four dimensions.

Table 3

Internal Consistency Coefficients for the Questionnaire Items (N = 92)

Dimension	Item	Pearson correlation with the dimension score
Dimension 1	A1	0.622**
	A2	0.752**
	A3	0.623**
	A4	0.639**
	A5	0.571**
	A6	0.446**

Table 3 to be continued

	A7	0.421**
	A8	0.356**
Dimension 2	A9	0.554**
	A10	0.713**
	A11	0.688**
	A12	0.737**
	A13	0.607**
	A14	0.543**
Dimension 3	A15	0.646**
	A16	0.684**
	A17	0.484**
	A18	0.516**
	A19	0.408**
	A20	0.473**
Dimension 4	A21	0.449**
	A22	0.723**
	A23	0.574**
	A24	0.615**

Note. ** $p < 0.01$.

Reliability. Reliability of the questionnaire was examined using Cronbach's alpha and split-half reliability corrected by the Spearman-Brown coefficient. As shown in Table 4, the values indicate acceptable to good reliability for all four dimensions of the questionnaire.

Table 4

Reliability Coefficients for the Questionnaire Dimensions (N = 92)

Dimension	No. of items	Cronbach's alpha	Split-half reliability
Dimension 1	6	0.762	0.758
Dimension 2	6	0.703	0.704
Dimension 3	6	0.712	0.703
Dimension 4	6	0.720	0.710

Data Collection Procedures

The questionnaire was created using Google Forms and shared with participants through WhatsApp. Ethical considerations were carefully addressed by ensuring voluntary participation, allowing participants to withdraw at any time, explaining the purpose of the study, and assuring confidentiality of responses for research purposes only (Cohen et al., 2018).

The collected data were analyzed using the Statistical Package for the Social Sciences (SPSS), Version 28. Descriptive statistics, including frequencies, percentages, means, standard deviations, interpretive levels, and item ranks, were used to describe participants' responses, demographic characteristics, and the mean scores of each dimension.

Before conducting the inferential analyses, the data were screened for outliers, and no influential outliers were detected. Data were also tested for normality using histograms, normal Q-Q plots, skewness and kurtosis values, and the Kolmogorov-Smirnov and Shapiro-Wilk tests. Overall, the normality indicators were considered acceptable across the demographic subgroups. Homogeneity of variance was also examined using Levene's test.

Independent-samples t-tests and one-way analysis of variance (One-Way ANOVA) were used to examine differences between groups. When the assumption of homogeneity of variance was violated, appropriate corrected statistics such as unequal variance t-test results, Welch's ANOVA, and Games-Howell post hoc comparisons were applied. These analyses were conducted to examine differences related to gender, teaching level, and years of teaching experience.

Results

Descriptive Results for Teachers' General Perceptions of Washback Effects

The first research question stated: *How do teachers perceive the washback effects of standardized examinations in intermediate and secondary schools in Kuwait?*

To answer this question, frequencies, percentages, means, and standard deviations were calculated for the items of the first dimension, *Teachers' general perceptions of washback effects*. It's important to note that negatively worded items were reverse-coded prior to analysis. Accordingly, lower mean scores indicate weaker alignment with positive washback and, therefore, reflect stronger negative washback effects. The results are presented in Table 5.

Table 5

Frequencies, Percentages, Means, and Standard Deviations for the Items of the Dimension Teachers' General Perceptions of Washback Effects (N = 92)

Item	Strongly agree	Agree	Neutral	Disagree	Strongly disagree	Mean	SD	Level	Rank
1. Standardized examinations help develop students' listening skills through appropriate tasks.	10 (10.9%)	24 (26.1%)	23 (25.0%)	24 (26.1%)	11 (12.0%)	2.98	1.20	Neutral	4
2. Standardized examinations help develop students' speaking skills through appropriate tasks.	9 (9.8%)	35 (38.0%)	16 (17.4%)	18 (19.6%)	14 (15.2%)	3.08	1.26	Neutral	3
3. Standardized examinations help develop students' reading skills through appropriate tasks.	21 (22.8%)	49 (53.3%)	12 (13.0%)	8 (8.7%)	2 (2.2%)	3.86	0.94	Agree	1
4. Standardized examinations help develop students' writing skills through appropriate tasks.	20 (21.7%)	49 (53.3%)	12 (13.0%)	8 (8.7%)	3 (3.3%)	3.82	0.98	Agree	2
5. Standardized examinations make students focus on grades rather than developing their knowledge in the English language.	2 (2.2%)	5 (5.4%)	13 (14.1%)	32 (34.8%)	40 (43.5%)	1.88	0.99	Disagree	6
6. Standardized examinations make me focus more on helping students achieve higher exam scores rather than helping them overcome their language difficulties.	2 (2.2%)	15 (16.3%)	11 (12.0%)	41 (44.6%)	23 (25.0%)	2.26	1.08	Disagree	5
Dimension Mean, SD, and Rank						2.98	0.72	Neutral	2

Notes. Items 5 and 6 were reverse-coded before calculating frequencies, percentages, means, standard deviations, levels, and ranks. Therefore, all reported responses and descriptive statistics are presented in the scored direction of the scale, where higher scores indicate more positive perceptions of washback effects.

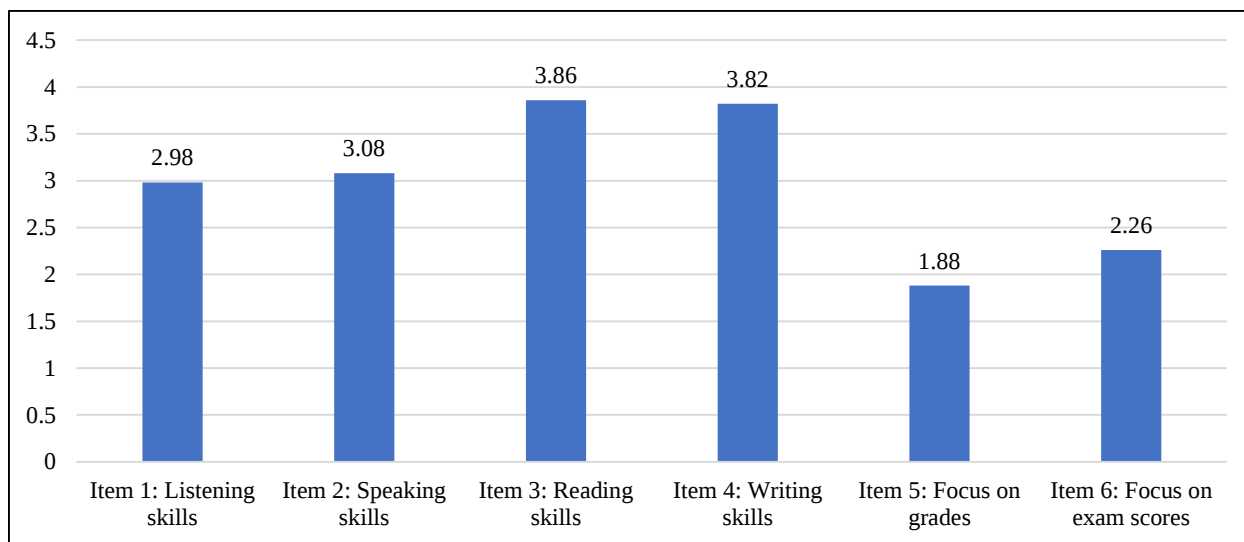


Figure 1. Mean scores of teachers' general perceptions of washback effects.

Item means ranged from (1.88) to (3.86). Item (3), "Standardized examinations help develop students' reading skills through appropriate tasks", ranked first with a mean of (3.86) and a standard deviation of (0.94), followed by Item (4), "Standardized examinations help develop students' writing skills through appropriate tasks", with a mean of (3.82) and a standard deviation of (0.98), and Item (2), "Standardized examinations help develop students' speaking skills through appropriate tasks", with a mean of (3.08) and a standard deviation of (1.26). These findings indicate that the highest positive perceptions were related to the perceived contribution of standardized examinations to the development of reading, writing, and, to a lesser extent, speaking skills.

In contrast, after reverse-coding, the lowest scored items were Item (5), "Standardized examinations make students focus on grades rather than developing their knowledge in the English language", with a mean of (1.88) and a standard deviation of (0.99), followed by Item (6), "Standardized examinations make me focus more on helping students achieve higher exam scores rather than helping them overcome their language difficulties", with a mean of (2.26) and a standard deviation of (1.08), and Item (1), "Standardized examinations help develop students' listening skills through appropriate tasks", with a mean of (2.98) and a standard deviation of (1.20). Therefore, the low means of Items (5) and (6) suggest that teachers perceived notable exam-oriented negative washback, particularly in terms of students focusing on grades and teachers prioritizing exam performance over addressing learners' language difficulties. Similarly, the relatively low mean of Item (1) indicates a weaker perception of the positive contribution of standardized examinations to listening skill development.

Overall, the highest positive perceptions were associated with reading- and writing-related effects. In contrast, the lowest-scoring responses point to stronger negative washback effects, as well as limited perceived support for listening skill development. There were no statistically significant differences in teachers' general perceptions of washback effects according to gender, teaching level, or years of experience.

Descriptive Results for the Influence of Standardized Examinations on Teaching Content

The second research question stated: *How do standardized examinations influence teaching content in intermediate and secondary schools in Kuwait?*

To answer this question, frequencies, percentages, means, and standard deviations were calculated for the items of the second dimension, *the influence of standardized examinations on teaching content*. Negatively

worded items were reverse-coded prior to analysis. Accordingly, lower mean scores indicate weaker alignment with positive washback and, therefore, reflect stronger negative washback effects. The results are presented in Table 6.

Table 6

Frequencies, Percentages, Means, and Standard Deviations for the Items of the Dimension the Influence of Standardized Examinations on Teaching Content (N = 92)

Item	Strongly agree	Agree	Neutral	Disagree	Strongly disagree	Mean	SD	Level	Rank
7. Standardized examinations align with curriculum objectives.	8 (8.7%)	45 (48.9%)	23 (25.0%)	10 (10.9%)	6 (6.5%)	3.42	1.02	Agree	2
8. Standardized examinations help me choose suitable teaching content to develop students' knowledge.	13 (14.1%)	38 (41.3%)	25 (27.2%)	16 (17.4%)	0 (0.0%)	3.52	0.94	Agree	1
9. My students expect me to teach them what is more likely to come in the exam.	0 (0.0%)	3 (3.3%)	10 (10.9%)	36 (39.1%)	43 (46.7%)	1.71	0.79	Strongly disagree	6
10. I give more time to content that will be tested and less time to content that will not be tested.	5 (5.4%)	11 (12.0%)	20 (21.7%)	33 (35.9%)	23 (25.0%)	2.37	1.15	Disagree	4
11. I neglect certain passages/parts from the textbook because I consider them less important for the examination.	10 (10.9%)	22 (23.9%)	12 (13.0%)	29 (31.5%)	19 (20.7%)	2.73	1.33	Neutral	3
12. I focus on certain tasks because I think similar ones will appear in the standardized examination.	4 (4.3%)	6 (6.5%)	14 (15.2%)	42 (45.7%)	26 (28.3%)	2.13	1.04	Disagree	5
Dimension Mean, SD, and Rank						2.65	0.62	Neutral	4

Notes. Items 9, 10, 11, and 12 were reverse-coded before calculating frequencies, percentages, means, standard deviations, levels, and ranks. Therefore, all reported responses and descriptive statistics are presented in the scored direction of the scale, where higher scores indicate more positive perceptions of washback effects.

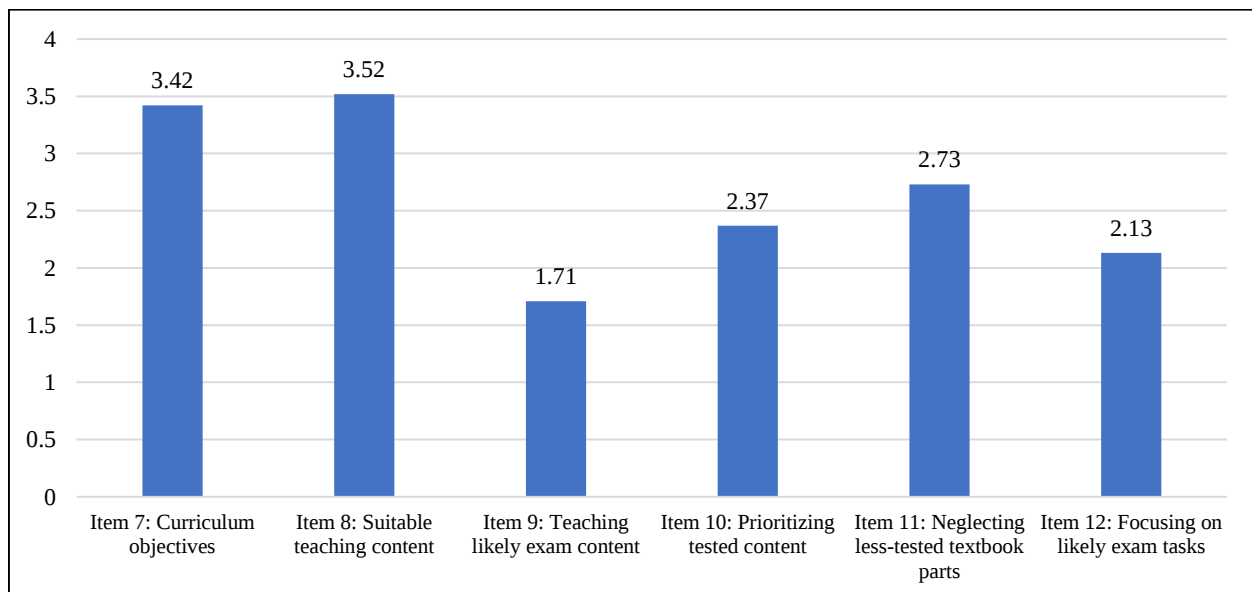


Figure 2. Mean scores for the influence of standardized examinations on teaching content.

Item means ranged from (1.71) to (3.52). Item (8), “Standardized examinations help me choose suitable teaching content to develop students’ knowledge”, ranked first with a mean of (3.52) and a standard deviation of (0.94), followed by Item (7), “Standardized examinations align with curriculum objectives”, with a mean of (3.42) and a standard deviation of (1.02), and, after reverse-coding, Item (11), “I neglect certain passages/parts from the textbook because I consider them less important for the examination”, with a mean of (2.73) and a standard deviation of (1.33). These findings indicate that the highest positive perceptions, reflecting positive washback, were related to the role of standardized examinations in helping teachers choose suitable teaching content and aligning teaching content with curriculum objectives.

In contrast, after reverse-coding, the lowest scored items were Item (9), “My students expect me to teach them what is more likely to come in the exam”, with a mean of (1.71) and a standard deviation of (0.79), followed by Item (12), “I focus on certain tasks because I think similar ones will appear in the standardized examination”, with a mean of (2.13) and a standard deviation of (1.04), and Item (10), “I give more time to content that will be tested and less time to content that will not be tested”, with a mean of (2.37) and a standard deviation of (1.15). The low mean scores for these negatively worded items indicate stronger negative washback related to students’ expectations, teachers’ focus on likely examination tasks, and the prioritization of tested content.

Overall, the highest positive perceptions, indicating positive washback, were associated with choosing suitable teaching content and alignment with curriculum objectives, whereas the lowest scored responses indicated stronger negative washback related to teaching content selection.

Differences in the influence of standardized examinations on teaching content according to demographic variables. To examine whether teachers’ perceptions of the influence of standardized examinations on teaching content differed according to selected demographic variables, independent-samples t-tests were used for gender and teaching level, whereas one-way ANOVA was used for years of experience. The results are presented in Table 7.

Table 7

Differences in the Influence of Standardized Examinations on Teaching Content According to Demographic Variables

Variable	N	M	SD	Test value	Df	P	Effect size
Gender							
Male	35	2.719	0.595	t = 0.875	90	0.384	d = 0.188
Female	57	2.602	0.636				
Teaching level							
Intermediate	65	2.736	0.617	t = 2.184	90	0.032	d = 0.500
Secondary	27	2.432	0.583				
Years of experience							
0-5 years	26	2.622	0.669	F = 1.765	2, 89	0.177	$\eta^2 = 0.038$
6-10 years	23	2.464	0.676				
More than 10 years	43	2.760	0.543				

Notes. M = Mean; SD = Standard Deviation; d = Cohen’s d; η^2 = eta squared.

The results in Table 7 show that there were no statistically significant differences in teachers’ perceptions of the influence of standardized examinations on teaching content according to gender or years of experience, whereas statistically significant differences were found according to teaching level. With regard to gender, male teachers reported a slightly higher mean score (M = 2.719, SD = 0.595) than female teachers (M = 2.602, SD = 0.636); however, this difference was not statistically significant (t = 0.875, p = 0.384).

For teaching level, the results showed a statistically significant difference ($t = 2.184$, $p = 0.032$), with intermediate school teachers obtaining a higher mean score ($M = 2.736$, $SD = 0.617$) than secondary school teachers ($M = 2.432$, $SD = 0.583$). This indicates that intermediate school teachers reported more positive perceptions of the influence of standardized examinations on teaching content compared with secondary school teachers. For years of experience, the results indicated no statistically significant differences among the three groups ($F = 1.765$, $p = 0.177$). Although teachers with more than 10 years of experience recorded the highest mean score ($M = 2.760$, $SD = 0.543$), followed by those with 0-5 years of experience ($M = 2.622$, $SD = 0.669$), and then those with 6-10 years of experience ($M = 2.464$, $SD = 0.676$), these differences were not significant.

Descriptive Results for the Influence of Standardized Examinations on Teaching Methods

The third research question stated: *How do standardized examinations influence teaching methods in intermediate and secondary schools in Kuwait?*

To answer this question, frequencies, percentages, means, and standard deviations were calculated for the items of the third dimension, *the influence of standardized examinations on teaching methods*. Negatively worded items were reverse-coded prior to analysis. Accordingly, lower mean scores indicate weaker alignment with positive washback and, therefore, reflect stronger negative washback effects. The results are presented in Table 8.

Table 8

Frequencies, Percentages, Means, and Standard Deviations for the Items of the Dimension the Influence of Standardized Examinations on Teaching Methods (N = 92)

Item	Strongly agree	Agree	Neutral	Disagree	Strongly disagree	Mean	SD	Level	Rank
13. Standardized examinations help me use different teaching methods to develop students' knowledge.	15 (16.3%)	35 (38.0%)	21 (22.8%)	14 (15.2%)	7 (7.6%)	3.40	1.16	Agree	3
14. Standardized examinations help me identify students' learning gaps and adjust my teaching methods accordingly.	14 (15.2%)	47 (51.1%)	14 (15.2%)	14 (15.2%)	3 (3.3%)	3.60	1.03	Agree	1
15. My principal/head of department/supervisor asks me to adopt teaching methods that ensure that students will get good marks on the exam.	0 (0.0%)	8 (8.7%)	25 (27.2%)	42 (45.7%)	17 (18.5%)	2.26	0.86	Disagree	6
16. I highly focus on drill and repetition methods because of the examination format.	2 (2.2%)	9 (9.8%)	26 (28.3%)	41 (44.6%)	14 (15.2%)	2.39	0.94	Disagree	5
17. I do not use the Communicative Language Teaching Method despite its importance because communicative skills are not part of the standardized examinations.	18 (19.6%)	35 (38.0%)	23 (25.0%)	11 (12.0%)	5 (5.4%)	3.54	1.10	Agree	2
18. I spend more time teaching test-taking techniques than language skills.	8 (8.7%)	24 (26.1%)	23 (25.0%)	29 (31.5%)	8 (8.7%)	2.95	1.13	Neutral	4
Dimension Mean, SD, and Rank						3.02	0.63	Neutral	1

Notes. Items 15, 16, 17, and 18 were reverse-coded before calculating frequencies, percentages, means, standard deviations, levels, and ranks. Therefore, all reported responses and descriptive statistics are presented in the scored direction of the scale, where higher scores indicate more positive perceptions of washback effects.

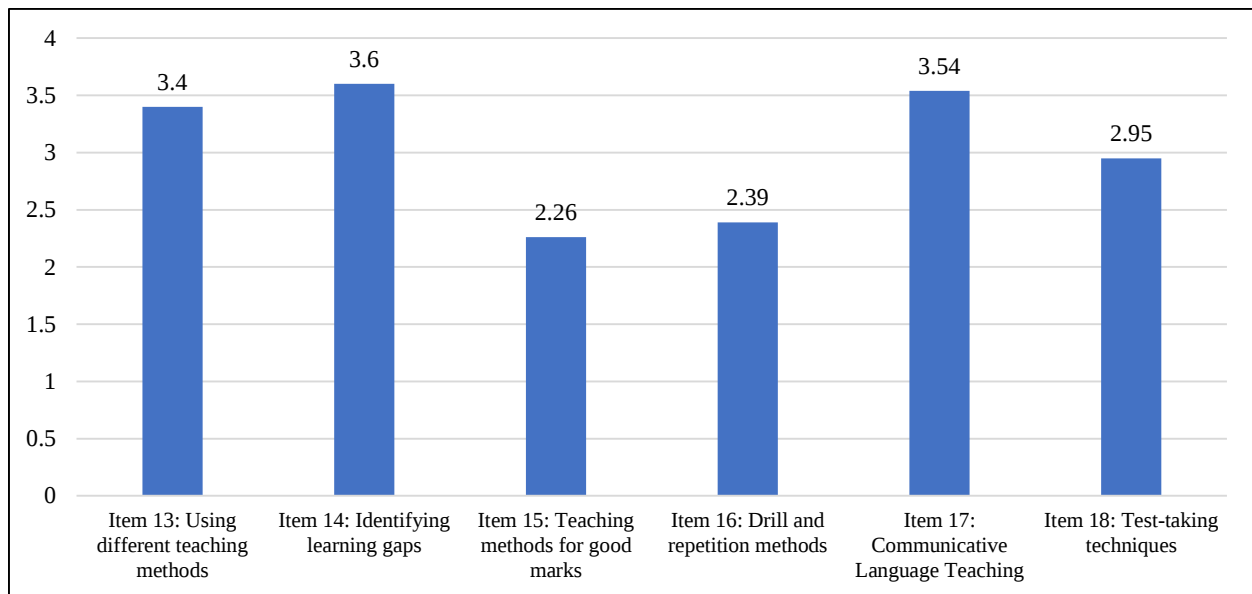


Figure 3. Mean scores for the influence of standardized examinations on teaching methods.

Item means ranged from (2.26) to (3.60). Item (14), “Standardized examinations help me identify students’ learning gaps and adjust my teaching methods accordingly”, ranked first with a mean of (3.60) and a standard deviation of (1.03), followed, after reverse-coding, by Item (17), “I do not use the Communicative Language Teaching Method despite its importance because communicative skills are not part of the standardized examinations”, with a mean of (3.54) and a standard deviation of (1.10), and Item (13), “Standardized examinations help me use different teaching methods to develop students’ knowledge”, with a mean of (3.40) and a standard deviation of (1.16). These findings indicate that the highest positive perceptions were related to using standardized examinations to identify students’ learning gaps and adjust teaching methods, maintaining the use of communicative language teaching despite examination constraints, and using different teaching methods to develop students’ knowledge.

In contrast, after reverse-coding, the lowest scored items were Item (15), “My principal/head of department/supervisor asks me to adopt teaching methods that ensure that students will get good marks on the exam,” with a mean of (2.26) and a standard deviation of (0.86), followed by Item (16), “I highly focus on drill and repetition methods because of the examination format”, with a mean of (2.39) and a standard deviation of (0.94), and Item (18), “I spend more time teaching test-taking techniques than language skills”, with a mean of (2.95) and a standard deviation of (1.13). The low mean scores for these negatively worded items indicate stronger negative washback related to administrative expectations, drill and repetition methods, and the use of test-taking techniques.

Overall, the highest positive perceptions, indicating positive washback, were associated with identifying learning gaps and adjusting teaching methods. In contrast, the lowest scored responses reflected stronger negative washback related to administrative expectations, drill and repetition methods, and the use of test-taking techniques. There were no statistically significant differences in teachers’ perceptions of the influence of standardized examinations on teaching methods according to gender, teaching level, or years of experience.

Descriptive Results for the Influence of Standardized Examinations on Classroom Assessment Practices

The fourth research question stated: *How do standardized examinations influence classroom assessment practices in intermediate and secondary schools in Kuwait?*

To answer this question, frequencies, percentages, means, and standard deviations were calculated for the items of the fourth dimension, *the influence of standardized examinations on classroom assessment practices*. Negatively worded items were reverse-coded prior to analysis. Accordingly, lower mean scores indicate weaker alignment with positive washback and, therefore, reflect stronger negative washback effects. The results are presented in Table 9.

Table 9

Frequencies, Percentages, Means, and Standard Deviations for the Items of the Dimension the Influence of Standardized Examinations on Classroom Assessment Practices (N = 92)

Item	Strongly agree	Agree	Neutral	Disagree	Strongly disagree	Mean	SD	Level	Rank
19. Standardized examinations have a positive effect on classroom assessment.	8 (8.7%)	33 (35.9%)	25 (27.2%)	19 (20.7%)	7 (7.6%)	3.17	1.10	Neutral	1
20. Standardized examinations allow me to focus on all four language skills while assessing students.	6 (6.5%)	27 (29.3%)	21 (22.8%)	26 (28.3%)	12 (13.0%)	2.88	1.17	Neutral	3
21. I rarely assess skills that are not included in the standardized examination.	3 (3.3%)	24 (26.1%)	31 (33.7%)	29 (31.5%)	5 (5.4%)	2.90	0.96	Neutral	2
22. Standardized examinations limit the kind of assessments I can use in my classroom.	2 (2.2%)	12 (13.0%)	24 (26.1%)	43 (46.7%)	11 (12.0%)	2.47	0.94	Disagree	4
23. I tend to copy standardized examination formats in my classroom tests to prepare the students for the exam.	2 (2.2%)	8 (8.7%)	16 (17.4%)	42 (45.7%)	24 (26.1%)	2.15	0.98	Disagree	6
24. Classroom assessment tends to emphasize standardized exam readiness over students' skill development.	4 (4.3%)	9 (9.8%)	25 (27.2%)	37 (40.2%)	17 (18.5%)	2.41	1.04	Disagree	5
Dimension Mean, SD, and Rank						2.66	0.58	Neutral	3

Notes. Items 21, 22, 23, and 24 were reverse-coded before calculating frequencies, percentages, means, standard deviations, levels, and ranks. Therefore, all reported responses and descriptive statistics are presented in the scored direction of the scale, where higher scores indicate more positive perceptions of washback effects.

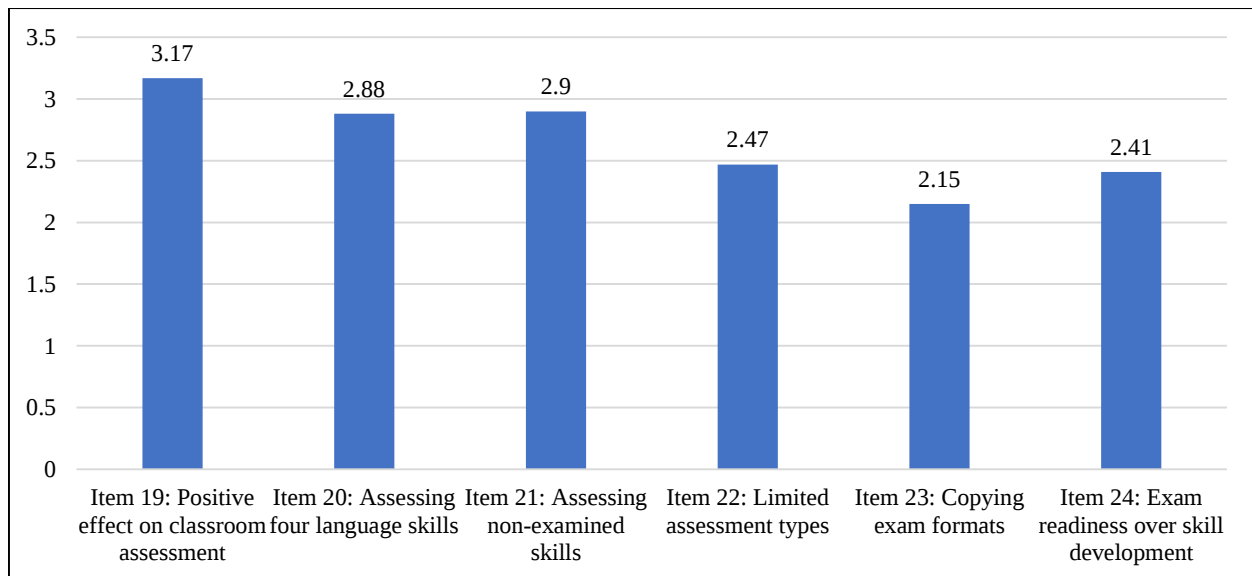


Figure 4. Mean scores for the influence of standardized examinations on classroom assessment practices.

Item means ranged from (2.15) to (3.17). Item (19), “Standardized examinations have a positive effect on classroom assessment”, ranked first with a mean of (3.17) and a standard deviation of (1.10), followed, after

reverse-coding, by Item (21), "I rarely assess skills that are not included in the standardized examination", with a mean of (2.90) and a standard deviation of (0.96), and Item (20), "Standardized examinations allow me to focus on all four language skills while assessing students", with a mean of (2.88) and a standard deviation of (1.17). These findings indicate that the highest scored responses were related to the perceived positive effect of standardized examinations on classroom assessment, teachers' tendency not to restrict classroom assessment to skills directly included in standardized examinations after reverse-coding Item (21), and the focus on the four language skills during classroom assessment.

In contrast, after reverse-coding, the lowest scored items were Item (23), "I tend to copy standardized examination formats in my classroom tests to prepare the students for the exam", with a mean of (2.15) and a standard deviation of (0.98), followed by Item (24), "Classroom assessment tends to emphasize standardized exam readiness over students' skill development", with a mean of (2.41) and a standard deviation of (1.04), and Item (22), "Standardized examinations limit the kind of assessments I can use in my classroom", with a mean of (2.47) and a standard deviation of (0.94). The low scores for these negatively worded items indicate stronger negative washback related to copying standardized examination formats, emphasizing examination readiness, and limiting the types of classroom assessment used.

Overall, the highest scored responses, indicating positive washback, were associated with the perceived positive effect of standardized examinations on classroom assessment and broader assessment of language skills, whereas the lowest scored responses indicated stronger negative washback related to copying assessment format, examination readiness, and limitations on classroom assessment practices. There were no statistically significant differences in teachers' perceptions of the influence of standardized examinations on classroom assessment practices according to gender, teaching level, or years of experience.

The results presented in Tables 5, 6, 8, and 9 indicate that all dimensions of the study fell within the Neutral level. The overall mean scores ranged from (2.65) to (3.02), suggesting that teachers' perceptions of the washback effects of standardized examinations in intermediate and secondary schools in Kuwait were moderate overall across general perceptions, teaching content, teaching methods, and classroom assessment practices.

Discussion

Teachers' General Perceptions of Washback Effects

The findings of teachers' general perceptions of the washback effects of standardized examinations showed the strongest positive perceptions regarding the development of reading and writing skills and, to a lesser extent, speaking skills. In contrast, teachers reported weaker positive perceptions regarding the development of listening skills. Although positive washback appeared to exist in relation to reading and writing, it was less noticeable in speaking and listening. This suggests that the washback effects of standardized examinations differed across language skills, with some skills receiving more positive washback than others. Similar patterns were reported in several previous studies, which showed that certain language skills tended to be neglected when they were not included in examinations. For example, Wall and Alderson (1996) reported an increased emphasis on reading skills, reduced attention to oral skills, and the exclusion of listening lessons because listening was not included in the examination (Wall & Alderson, 1996). Likewise, Al-Amin and Greenwood (2018) found that many teachers recognized the importance of listening and speaking for developing English proficiency. However, they argued that examinations needed to include these skills before they could become part of classroom practice (Al Amin & Greenwood, 2018). Rahman et al. (2024) also pointed out that neglecting listening and speaking skills

in assessment and testing systems may lead to weak communication abilities among learners. Therefore, it may be suggested that the weaker positive washback perceptions related to speaking and listening were associated with the absence of these skills from the standardized examinations used in intermediate and secondary schools in Kuwait.

The findings also revealed noticeable exam-oriented negative washback. Teachers perceived that students were more concerned with grades, while teachers focused more on students' examination performance than on addressing their language difficulties. The strong focus on grades and examination performance is consistent with the findings of Ali and Hamid (2020), which suggested that the social importance attached to grades and the pressure to achieve high scores in public examinations may negatively influence teaching and learning. Ali and Hamid (2020) reported that one teacher participant stated that "now-a-days, people in our country value grades more than what students learn" (p. 137). The participant also expressed concern that teachers are held accountable when students do not achieve passing grades or high examination scores. Similarly, Choi (2008) pointed out that the widespread importance of English led many learners to prioritize achieving high EFL examination scores rather than focusing on the actual goals of language learning. In addition, the findings are consistent with Al-Wadi's (2020) study, which found that examinations restricted teachers' ability to address students' individual learning needs. Likewise, Al-Hinai (2023) found that teachers focused more on examination performance, while Ministry of Education decision-makers emphasized broader communicative outcomes, creating inconsistencies in implementation, particularly in speaking and listening assessment. Das et al. (2014) further suggested that this misalignment creates negative washback by encouraging teachers to prioritize examination preparation over meaningful learning. Therefore, it may be suggested that standardized examinations create negative washback by encouraging teachers to place greater emphasis on students' examination scores than on addressing their language difficulties, while also encouraging students to prioritize grades over developing their English language knowledge.

The Influence of Standardized Examinations on Teaching Content

The findings related to the influence of standardized examinations on teaching content revealed both positive and negative washback effects. Positive washback was related to the role of standardized examinations in helping teachers select suitable teaching content and align teaching content with curriculum objectives. This finding is consistent with Cheng's (1997) study, in which many teachers reported that the revised HKCEE examination encouraged the integration of language skills and promoted more interactive classroom activities through discussion and participation. However, regarding the alignment between teaching content and curriculum objectives, the findings contradict those of Al-Hinai (2023), who identified a mismatch between the examination and curriculum implementation that created a gap between intended goals and actual classroom practice. Similarly, Al-Amin and Greenwood (2018) argued that the national examination system in Bangladesh secondary education made the achievement of curriculum goals and objectives impractical. This may suggest that positive washback is more likely when examinations are aligned with curriculum objectives. The findings also showed that many teachers disagreed with the statement, "I neglect certain passages/parts from the textbook because I consider them less important for the examination", but still believed that negative washback existed in terms of giving more time to tested content and less time to content that would not appear in the examination. This may suggest that teachers do not completely neglect parts of the textbook, but they still prioritize tested content. Similar findings were reported by Al-Amin and Greenwood (2018), who found that some teachers attempted to

teach all chapters in the required textbooks, although they focused more on important examination content as the examination approached. Likewise, Wall and Alderson (1996) found that teachers generally attempted to cover the textbooks, which was considered positive, although it became negative when it involved memorizing texts likely to appear in examinations (Wall & Alderson, 1996).

The findings also indicated stronger negative washback related to teachers focusing on certain tasks because they expected similar tasks to appear in the standardized examination. This is consistent with Li and Hu's (2025) findings, which showed that teachers focused more on vocabulary and speaking practice, while writing received the least attention because it was less emphasized in the examination. Aron (2015) also found that teachers focused heavily on practicing speaking test formats. Although these activities may help students prepare for examinations, they may be less engaging than presentations or class discussions that teachers might use in the absence of examinations (Aron, 2015). Similarly, Al-Hinai (2023) found that although the new English elective diploma examination in Oman aimed to promote positive change, it also produced partly negative washback because teachers focused narrowly on examination-related tasks instead of broader language development. The findings also revealed stronger negative washback related to students expecting teachers to focus on content likely to appear in the examination. This is consistent with Ali and Hamid's (2020) findings. One teacher participant explained that she mainly selected topics that were important for the examination because this was what her students expected (Ali & Hamid, 2020). She also reported that students opposed attempts to focus on skills that were not tested (Ali & Hamid, 2020). These findings may suggest that standardized examinations create negative washback by encouraging a narrow focus on examination task types and increasing students' expectations that teachers should focus mainly on content likely to appear in examinations.

The Influence of Standardized Examinations on Teaching Methods

The findings regarding the influence of standardized examinations on teaching methods showed that the strongest positive perceptions were associated with helping teachers recognize students' learning gaps and modify their teaching methods based on students' needs. This is consistent with Al-Wadi (2020), whose findings suggested that teachers viewed the national examination as an effective tool that directed teaching toward clear and specific purposes. However, although teachers believed that examinations helped them focus on students' immediate learning needs, the examinations were considered less effective in supporting the development of new learning experiences (Al-Wadi, 2020). The findings also showed positive perceptions regarding continuing to use communicative language teaching despite examination constraints. This contradicts the finding of Li and Hu (2025), who found that lecture-based teaching was used most frequently, while communicative activities were used less often because of time limitations and examination preparation demands. Despite these constraints, teachers became more attentive to students and more rigorous in their teaching (Li & Hu, 2025). This supports the finding that standardized examinations may help teachers identify students' learning needs and adapt their teaching methods. The findings also revealed positive perceptions regarding the use of a variety of teaching methods to improve students' knowledge, indicating relatively positive washback on teaching methods. This differs from the findings of Al-Wadi (2020), which reported that national examinations limited teachers from trying new teaching methods. Overall, the findings suggest that teachers did not believe that standardized examinations in Kuwait negatively influenced their use of communicative language teaching or different teaching methods to support students' learning.

In contrast, the findings regarding the influence of standardized examinations on teaching methods revealed stronger negative washback related to administrative expectations that teachers should use teaching methods aimed at helping students achieve high examination scores. This is consistent with the findings of Al Amin and Greenwood (2018), who reported that principals and headmasters asked teachers to adopt teaching methods that would ensure good examination marks. The findings also showed stronger negative washback related to the use of drill and repetition methods because of the examination format. This aligns with Liang et al. (2025), who found that repetitive practice papers and test-taking strategies occupied much of the classroom time because of examination demands. Similarly, the findings showed that teachers spent more time teaching test-taking techniques than developing language skills. This corresponds with the findings of Aron (2015), who reported that teachers frequently focused on explaining examination formats, giving special instructions and hints about examinations, and informing students how to achieve higher scores. Likewise, Li and Hu (2025) reported that teachers regularly included test-related explanations and strategies in their lessons and organized instruction around examination requirements. Overall, the findings suggest that standardized examinations may encourage teaching practices that focus more on examination preparation and test-taking strategies than on broader language skill development.

The Influence of Standardized Examinations on Classroom Assessment Practices

The findings regarding the influence of standardized examinations on classroom assessment practices indicated some positive washback on classroom assessment. This aligns with the findings of Al-Wadi (2020), who found that teachers viewed the national examination as having a positive influence on assessment practices by guiding the development of classroom tests and helping them determine the focus of feedback provided to students. Teachers also considered the examination a stimulus for reflecting on their teaching and assessment practices to better support students' EFL/ESL learning (Al-Wadi, 2020). The findings further revealed positive washback related to teachers' tendency not to limit classroom assessment to skills directly included in standardized examinations and to maintain attention to all four language skills during classroom assessment. This differs from the findings of Azadi and Gholami (2013), who found that the final examination did not necessarily ensure the learning of all required skills, such as listening, speaking, and writing, and that classroom tests were strongly influenced by the direction of the final examination. Overall, the findings suggest that standardized examinations may support classroom assessment practices without restricting teachers' attention to all four language skills or limiting assessment to tested skills only.

In contrast, the findings also revealed some negative washback on classroom assessment practices. The strongest negative washback was related to copying standardized examination formats. This is in line with Aron (2015), who found that five of the seven schools in the study organized mock examinations using previous EFL examination papers. Similarly, Wall and Alderson (1996) reported negative washback associated with the direct use of materials from past examination papers. Negative washback was also evident in the emphasis on examination readiness. This corresponds with the findings of Al Amin and Greenwood (2018), who argued that the rote learning encouraged by examinations narrows both perceptions of learning and learning experiences, reducing language learning to the acquisition of correct answers rather than the development of communication and thinking skills. The findings further showed that teachers perceived standardized examinations as limiting the types of assessments used in their classrooms. This aligns with Al-Wadi's (2020) findings, which indicated that teachers believed the national examination limited students' learning experiences to the knowledge and skills

assessed by the examination. As a result, students were less likely to be exposed to other assessment experiences that could evaluate their performance in different ways (Al-Wadi, 2020). Overall, the findings suggest negative washback related to copying examination formats, emphasizing examination readiness, and limiting the types of classroom assessment methods.

Overall Interpretation of Findings

The findings of this study support previous washback research (e.g., Al-Hinai, 2023; Li & Hu, 2025; Liang et al., 2025), highlighting that washback is a complex, indirect, multidirectional, and context-dependent phenomenon. As noted by Aron (2015), it is often difficult to determine whether a particular washback effect is entirely positive or negative because the same teaching practice may have both beneficial and harmful consequences. This suggests that washback should not be viewed as a simple positive–negative dichotomy, but rather as a phenomenon whose effects depend on the specific educational context. The findings showed that the strongest positive perceptions were related to the role of standardized examinations in developing reading and writing skills. This is understandable in the Kuwaiti context, where standardized examinations at the intermediate and secondary levels mainly assess reading and writing. Consequently, weaker perceptions regarding speaking and listening may be related to the absence of these skills from the examinations.

In light of these findings, standardized examinations may not fully align with curriculum objectives, which emphasize the integration of the four language skills and communicative learning. Nevertheless, teachers reported that standardized examinations help them align instruction with curriculum objectives and select appropriate teaching content. Findings related to teaching methods and classroom assessment may help explain this pattern. Teachers reported continuing to use communicative language teaching despite examination constraints and maintaining attention to all four language skills during classroom assessment. They also perceived standardized examinations as helping them identify students' learning gaps, adapt and diversify their teaching methods, and improve classroom assessment practices. These findings suggest that standardized examinations may serve as a guide for teachers, enabling them to address what is missing from the examinations through their own teaching and classroom assessment practices.

Furthermore, teachers also reported that standardized examinations limit the range of assessment methods used in their classrooms. This does not appear to mean that teachers neglect untested content or language skills. Rather, it may reflect a greater emphasis on tested content, examination formats, examination tasks, and examination readiness. Therefore, teachers may continue to assess all four language skills while giving priority to examination-related requirements.

The findings also revealed several forms of negative washback on teachers. Teachers perceived that standardized examinations encourage greater attention to students' examination scores than to their language learning difficulties. Negative washback was also reflected in a stronger focus on examination task types, the use of drill and repetition methods, increased emphasis on test-taking strategies, and the copying of standardized examination formats in classroom assessment. Together, these findings indicate that standardized examinations influence teaching content, teaching methods, and classroom assessment by directing greater attention toward examination preparation. In addition, teachers perceived negative washback related to students prioritizing grades and examination content over developing their English language abilities. They also perceived negative washback related to administrators, who were expected to encourage practices that help students achieve higher examination scores.

Conclusion

The findings of this study indicate that standardized examinations in Kuwait generate both positive and negative washback effects on teaching and learning, suggesting that washback is a complex phenomenon. Positive washback was reflected in teachers' perceptions that standardized examinations contribute to the development of reading and writing skills and, to a lesser extent, speaking skills. In contrast, teachers reported weaker positive perceptions regarding the development of listening skills. This may be related to the fact that standardized examinations at the intermediate and secondary levels in Kuwait mainly assess reading and writing. The findings also revealed negative washback related to teachers placing greater emphasis on students' examination scores than on their language learning difficulties, as well as students prioritizing grades over developing their English language abilities.

The findings revealed positive washback related to the role of standardized examinations in helping teachers select suitable teaching content and align instruction with curriculum objectives. The findings also suggested that teachers do not completely neglect untested content, but they tend to prioritize content that is more likely to appear in examinations. Negative washback was evident in the stronger focus on certain examination task types. The findings also showed that students expected teachers to concentrate mainly on exam-related content.

The findings regarding teaching methods showed positive washback related to helping teachers identify students' learning gaps, adjust and diversify their teaching methods according to students' needs, and continue using communicative language teaching despite examination constraints. These findings suggest that teachers did not believe that standardized examinations negatively influenced their use of communicative language teaching or different teaching methods to support students' learning. At the same time, negative washback was reflected in administrative expectations that teachers should use methods aimed at helping students achieve higher examination scores, the use of drill and repetition methods, and increased attention to test-taking techniques.

The findings regarding classroom assessment practices indicated some positive washback, particularly in teachers' tendency not to limit assessment to skills directly included in standardized examinations and to maintain attention to all four language skills during classroom assessment. These findings suggest that standardized examinations may support classroom assessment practices without restricting teachers' attention to all four language skills. However, the findings also revealed negative washback related to copying standardized examination formats, emphasizing examination readiness, and limiting the types of assessments used in classrooms. This does not appear to mean that teachers neglect untested content or language skills. Rather, teachers may continue to assess all four language skills while giving greater priority to exam-related requirements.

This study provides insights into how standardized examinations influence teaching content, teaching methods, and classroom assessment practices in Kuwaiti public schools. It may encourage teachers to reflect on how examination requirements influence their classroom practices and how they can balance exam preparation with broader language learning goals. The findings may help policymakers in Kuwait better understand the consequences of standardized examinations and support efforts to improve examination systems by reducing negative washback and strengthening positive washback. The study contributes to the limited body of research on washback in the Kuwaiti context and adds to the broader washback literature by examining teachers' perceptions across teaching content, teaching methods, and classroom assessment practices. Furthermore, it provides a foundation for future studies involving classroom observations, interviews, students, or other stakeholders to gain a more comprehensive understanding of washback effects.

The use of non-random sampling and the limited time available to include more participants or conduct interviews represent limitations of this study. As the findings are based on questionnaire data and rely on teachers' perceptions, further research is needed to determine whether these perceptions reflect actual classroom practices through classroom observations. Further research is also needed to investigate washback more comprehensively by conducting interviews and involving students and other stakeholders.

References

- Al Amin, M., & Greenwood, J. (2018). The examination system in Bangladesh and its impact: On curriculum, students, teachers and society. *Language Testing in Asia*, 8(4), 1-18. Retrieved from <https://doi.org/10.1186/s40468-018-0060-9>
- Alderson, J. C., & Wall, D. (1993). Does washback exist? *Applied Linguistics*, 14(2), 115-129. Retrieved from <https://doi.org/10.1093/applin/14.2.115>
- Al-Hinai, S. I. M. (2023). Investigating the washback of English language testing reform on teaching English in Oman (Doctoral dissertation, University of York, 2023).
- 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. Retrieved from <https://doi.org/10.1080/15434303.2020.1717495>
- Al-Wadi, H. M. (2020). Bahrain's secondary EFL teachers' beliefs of English language national examination: "How it made teaching different?" *International Journal of Instruction*, 13(1), 197-214. Retrieved from <https://doi.org/10.29333/iji.2020.13113a>
- Aron, K. (2015). Teachers' perceptions of the washback effect of the Year 12 English examination (Master's thesis, University of Tartu, 2015).
- Azadi, G., & Gholami, R. (2013). Feedback on washback of EFL tests on ELT in L2 classroom. *Theory and Practice in Language Studies*, 3(8), 1335-1341. Retrieved from <https://doi.org/10.4304/tpls.3.8.1335-1341>
- Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice: Designing and developing useful language tests*. Oxford: Oxford University Press.
- Brown, H. D., & Abeywickrama, P. (2019). *Language assessment: Principles and classroom practices* (3rd ed.). Harlow: Pearson.
- Cheng, L. (1997). How does washback influence teaching? Implications for Hong Kong. *Language Education*, 11(1), 38-54.
- Cheng, L., & Curtis, A. (2004). Washback or backwash: A review of the impact of testing on teaching and learning. In L. Cheng, Y. Watanabe, and A. Curtis (Eds.), *Washback in language testing: Research contexts and methods* (pp. 3-17). Mahwah: Lawrence Erlbaum Associates.
- Choi, I.-C. (2008). The impact of EFL testing on EFL education in Korea. *Language Testing*, 25(1), 39-62. Retrieved from <https://doi.org/10.1177/0265532207083744>
- Cohen, L., Manion, L., & Morrison, K. (2018). *Research methods in education* (8th ed.). New York: Routledge.
- Creswell, J. W. (2009). *Research design: Qualitative, quantitative, and mixed methods approaches* (3rd ed.). London: SAGE Publications.
- Das, S., Shaheen, R., Shrestha, P., Rahman, A., & Khan, R. (2014). Policy versus ground reality: Secondary English language assessment system in Bangladesh. *The Curriculum Journal*, 25(3), 326-343. Retrieved from <https://doi.org/10.1080/09585176.2014.909323>
- Huang, L., & Zeng, J. (2025). Sustainability of washback effects: A longitudinal study of China's twice-yearly NMET reform. *Language Testing in Asia*, 15, Article 19. Retrieved from <https://doi.org/10.1186/s40468-025-00358-9>
- Hughes, A. (2003). *Testing for language teachers* (2nd ed.). Cambridge: Cambridge University Press.
- Li, Y., & Hu, B. (2025). A study on the washback effect of HSK on international Chinese language education. *SAGE Open*, 15(3), 1-14. Retrieved from <https://doi.org/10.1177/21582440251363774>
- Liang, Y., Siraj, S., & Lo, Y. Y. (2025). Exploring washback processes in language teaching through teachers' beliefs in the Chinese National Matriculation English Test (NMET). *Humanities and Social Sciences Communications*, 12, Article 1127. Retrieved from <https://doi.org/10.1057/s41599-025-05345-2>
- Mansory, M. M. (2016). EFL teachers' beliefs and attitudes towards English language assessment in a Saudi University's English Language Institute (Doctoral dissertation, University of Exeter, 2016).
- McNeil, L. M. (2000). *Contradictions of school reform: Educational costs of standardized testing*. New York: Routledge.

- Ministry of Education. (2026). *Teacher's library*. Retrieved from <https://elibrary.moe.edu.kw/TeachersLibrary>
- Phelps, R. P. (2005). The rich, robust research literature on standardized testing's achievement benefits. In R. P. Phelps (Ed.), *Defending standardized testing* (pp. 55-90). Mahwah: Lawrence Erlbaum Associates.
- Rahman, M., Jamila, M., & Siam, M. R. (2025). Language skills development: EFL teachers' and learners' perceptions of the existing teaching-learning practices. *The International Journal of Literacies*, 32(1), 1-20. Retrieved from <https://doi.org/10.18848/2327-0136/CGP/v32i01/1-20>
- Sireci, S. G. (2005). The most frequently unasked questions about testing. In R. P. Phelps (Ed.), *Defending standardized testing* (pp. 111-121). Mahwah: Lawrence Erlbaum Associates.
- Tryzna, M. M., & Al Sharoufi, H. (2017). English language education policy in Kuwait. In R. Kirkpatrick (Ed.), *English language education policy in the Middle East and North Africa* (pp. 77-91). New York: Springer.
- Wall, D., & Alderson, J. C. (1996). Examining washback: The Sri Lankan impact study. In A. Cumming and R. Berwick (Eds.), *Validation in language testing* (pp. 194-222). Bristol: Multilingual Matters.
- Xie, Q., & Jia, Q. (2025). Three decades of research on washback (1993-2023): A bibliometric study. *Language Testing in Asia*, 15(23), 1-29. Retrieved from <https://doi.org/10.1186/s40468-025-00357-w>