Teachers as Key Factors in Washback: analyzing classroom practices in Libyan preparatory schools

Hamuda, Mustafa

Department of English Language, University of Tripoli, Libya

mostafa008@hotmail.co.uk

Submitted: 16/01/2025 **Accepted**: 02/02/2025 **Published** 1/03/2025

ABSTRACT

This paper investigates the washback effects of a national achievement test on teaching practices in three preparatory schools located in the northwest Libya. Adopting a case-study methodology, this paper provides an analysis of how national testing influences classroom teaching. Data were collected through semi-structured interviews with teachers, and video recordings of classroom lessons. Inductive thematic analysis was used to examine the interview transcripts and the recorded classroom interactions. The study applies Alderson and Wall's (1993) washback model, along with the objectives outlined in the course book materials, as benchmarks to analyse classroom practices. The findings reveal that contextual factors, particularly the teachers' own characteristics, play a critical role in shaping their teaching.

Keywords: Washback, Communicative Approach, Contextual Factors, Teachers' Practices, National Achievement Test.

الملخص

تركز هده الدراسة على تأثير الامتحان النهائي على العملية التدريسية في مرحلة إنهاء التعليم الأساسي في ثلاث مدارس في شمال غرب ليبيا. تم اتباع الأثنوجرافيا ودراسة الحالة في تصميم الدراسة. تم جمع البيانات من المدرسين بالمقابلات الشخصية والتسجيل المصور للحصص. استخدم الباحث طريقة التحليل الاستقرائي لتحليل الفكرة للمقابلات والحصص. نم استخدام نموذج الدارسون ووال (1993)، الامتحان انهائي وكدلك أهداف المنهج كإطار لتحليل المناقشات والدروس المسجلة. وقد اثبت الدراسة ان عامل المدرسين كجزء من العوامل يحدد التدريس وكيفيته.

الكلمات المفتاحية: تأثير الامتحان النهائي، الطريقة التفاعلية، العوامل السياقية، ممارسات المدرسين، الامتحان النهائي.

Introduction

Language tests have always been the cornerstone in classroom teaching. Final achievement tests, however, are used to make important decisions which have direct effects on teachers. The relationship between language testing and teaching has long been a subject of interest in educational research, particularly in the context of high-stake tests. They are high-stakes because the stakes associated with the results of passing or failing the tests. In many educational systems, the influence of such tests is significant, affecting not only the curriculum but also the way teachers (and learners) approach their lessons, an area of research defined as washback.
Earlier washback research emphasized that high-stakes tests might determine the teaching process (e.g., Messick,1996; Alderson & Wall, 1993), a position some washback researchers in Libya still hold. It is beyond doubt that teachers naturally tend to focus on test demands, but attributing what happens in the classroom to tests only is questionable. In fact, washback research have shown that washback is shaped by contextual factors, and researchers need to study washback in their own contexts (Rea-Dickins & Scott, 2007; Alderson, 2004; Cheng & Curtis, 2004; Watanabe, 2004); washback is context-specific.

Understanding the nature of context within educational systems and the roles of teachers in those contexts is central in attempting to establish evidentially a link between high-stakes tests and their effects. This is big challenge washback researchers are still facing.

In Libya, a new curriculum based on the communicative approach was introduced into preparatory schools in the late 1990s, while a new version of the final school test consisting of multiple-choice and true/false questions was introduced in 2007. Where English is a mandatory subject at the preparatory school level, the role of the final test in shaping instructional methods and classroom dynamics is important. The final test in Libya's educational system is based on multiple-choice questions (MCQs) that focus on decontextualized grammar and vocabulary, a stark contrast to the communicative language teaching approach that is the foundation of the course materials. Investigating the washback effect of the new test poses challenges in untangling pre-test teaching practices, post-test changes, and related contextual factors.

Despite the centrality of high-stakes tests, there is limited research on the final test effects in the Libyan context, especially with respect to

how teachers' understanding of language teaching methodologies intersects with their test preparation practices. The study aims to explore

the relationship between high-stakes testing and classroom teaching practices in Libyan third preparatory schools. It investigates how the structure and content of the final test might influence teachers' instructional practices, their use of language in the classroom, and their perceptions of effective language teaching.

This study seeks to achieve the following objectives:

1. To examine how teachers in Libyan preparatory schools understand and implement the communicative approach prescribed in the course materials.
2. To explore the degree to which the final test influences teachers' instructional methods and classroom practices.
3. To analyse the alignment, or the lack of alignment, between the course materials, the communicative approach, and the format of the final test.
4. To identify the factors that mediate the influence of the final test on teachers' practices, including teachers' attitudes toward language teaching, and societal expectations.

To address these objectives, the study is guided by the following research questions:

1. How does the final test shape teachers' teaching methods and classroom activities?
2. How do contextual factors, such as teachers' knowledge, and societal expectations, influence their response to the final test?

This study aims to contribute to washback research by exploring Libyan preparatory education and providing insights for future educational policy. The findings are expected to provide a deeper understanding of the complexities surrounding the implementation of communicative teaching methods in relation to a high-stakes testing. The following two sections provide some information about the course materials taught, and the final test being administered.

The Course Materials

Previous textbooks, used before the current materials, introduced by the Ministry of Education, and designed by Mr. Mustafa Gusbi in the 1960s, were entitled 'English for Libya'. They focused on the learning of basic sentence patterns using direct method (Barton, 1968).This series proved ineffective as it led to an overemphasis on oral drills, correct grammar and pronunciation, and memorisation of isolated vocabulary and grammar (Shihiba, 2011; Orafi & Borg, 2009; Aldabbus, 2008; Alhmali, 2007). The new course materials were introduced in the 1990s and were designed to be taught from the first year of the preparatory schools (age11) until the final year at secondary school (age17).The content of

the course materials has been designed on the principles of the communicative approach, with a variety of tasks and activities which stimulate learners to engage in meaningful discussions through pair work, group work and whole class work .The teacher's book has been designed to guide teachers in how to teach each lesson in the course materials. Libyan students' study nine subjects in the third (final) year at preparatory school.
All the subjects are in Arabic except for that of English language. Students have to pass tests in all of the subjects in order to obtain their preparatory certificates.

Description of the Test

It is the first high-stakes achievement test in the English language students take, usually at the end of May. It is a high-stakes test because of its fundamental role in determining the learners' future study as it is the gateway to secondary school. It is designed to assess the students' achievement relative to the course material they have studied in their academic year. The actual test which is the subject of this analysis is an achievement test administered in Libyan preparatory schools as a trial version which is representative of the actual final test of the academic year 2008/2009. The test consisted of **thirty-six** questions, of which **eighteen** were true/false, **twelve** were multiple-choice, **five** are matching, and **one** was wh-question. It tests decontextualized vocabulary and grammar, and (memorisation of) course content. It was printed on four pages with double-line spacing which made it easy for the learners to read. No marking criteria were provided for each question which, consequently, would prevent the learners, and the teachers, from having a clear idea as to how the different test components could be marked. This might prevent them from allocating their time and attention in accordance with the marks allocated to test each item. Furthermore, the limited range of question-types; just multiple-choice and true-false was a potential weakness. It did not test the fundamental skills of listening, speaking and writing.

Literature Review

This section examines the key concepts related to washback in language testing and teaching, synthesizing findings from previous studies, particularly in the context of high-stakes tests. A key aim of this review is to identify the research gap in the existing literature, particularly within the Libyan context, and to position this study as a significant contribution to enhancing the understanding of washback effects in this setting.
While the foundational work on washback can be traced back to Spolsky (1981), who first addressed the high-stakes nature of language tests,

significant research emerged only in the mid-2000s. Much of the early washback research provided essential frameworks and insights that continue to inform contemporary studies. This review will focus primarily on the research conducted during this foundational period, as it contributed significantly to the development of washback theory.

Washback Definition

Washback refers to the influence that high-stakes tests exert on classroom practices. While the terms "washback" and "impact" are often used interchangeably, there is a critical distinction between them. "Impact" typically refers to broader effects on the education system and society, whereas washback specifically addresses its influence on classroom instruction (Wall, 2005). This distinction is important for understanding how tests shape teaching practices in direct, immediate ways. Bachman and Palmer (1996) further clarified this by describing the micro- and macro-level effects of tests, with washback focusing on the former—the way tests affect classroom practices.

Washback: Positive, Negative or Both?

The consequences of washback are often categorized as either positive or negative, though the bidirectional nature of this influence is not always straightforward. Some scholars, including Bailey (1996), Messick (1996) and Alderson and Wall (1993), have recognized that washback can both promote (positive) and inhibit (negative) teaching practices. Messick (1996) highlighted that tests might either foster or impede effective teaching. Spolsky (1981) argued that washback should primarily be viewed in terms of unintended negative side effects, drawing attention to how tests might inadvertently harm teaching practices, even when they aim to improve educational outcomes.

In contrast, Alderson and Hamp-Lyons (1996) and Cheng and Curtis (1993) expanded this view by suggesting that washback could be both intentional and unintentional, with effects that might be either positive or negative. This view is crucial for understanding the complexity of washback in educational settings. For example, tests can be leveraged to bring about beneficial changes in teaching, as Pearson (1988) suggested, using tests as 'levers for change' to drive educational reform. Popham (1987) introduced the concept of measurement-driven instruction (MDI), arguing that tests, when properly designed, would guide the curriculum toward desired educational outcomes. For positive effects, Hughes (2003) suggested the following:

-test the abilities you want to encourage.
-sample widely and unpredictably.
-use direct testing.
-make testing criterion-referenced.
-base achievement tests on objectives.

-ensure the test is known and understood by teachers.
-provide assistance to teachers. (pp.53-56)

This was an attempt to establish a framework for designing tests. Nevertheless, it would need deep theoretical knowledge about the right ability to be tested.

On the other hand, negative washback occurs when tests are designed with a narrow definition of language ability, which can limit the scope of teaching (Shohamy, 2001; Linn, 2000; Madaus, 1988;). When tests focus primarily on discrete elements such as grammar and vocabulary, teachers may feel compelled to narrow the curriculum and prioritize test preparation over more holistic teaching methods (Pedulla et al., 2003). This, in turn, can lead to teaching practices that focus on rote memorization rather than fostering deeper understanding and critical thinking (Cheng & Curtis, 2004). Spolsky (1981) argued that multiple-choice tests, in particular, would encourage a narrow focus on test preparation, which might reduce the overall quality of education. Similarly, Fredericksen (1984) noted that such tests tended to encourage teaching that prioritized factual knowledge over higher-order thinking skills.

Some researchers, such as Madaus (1988) and Popham (1987) pointed out that the relationship between washback and the stakes of a test would be more significant than its technical quality. In cases where schools are held accountable for their performance on high-stakes tests, these tests often become a tool for shaping what is taught in the classroom. This phenomenon, often referred to as 'teaching to the test', can lead to negative washback by reducing the breadth and depth of instruction (Shohamy, 2001).

However, whether high-stakes tests inevitably lead to positive or negative washback remains a debated issue. Taylor (2005) argued that it would be overly simplistic to assume a direct, one-to-one relationship between tests and teaching practices. Research has shown that washback is a more complex phenomenon, influenced by various contextual factors (Cheng, 2008; Rea-Dickins & Scott, 2007; Spratt, 2005; Alderson, 2004). Cumming (2009) and Wall and Alderson (1993) emphasized that washback could be mediated by a variety of factors, including psychological, social, cultural, and political elements. These factors differ across educational contexts, and their interaction can significantly shape the nature of washback. Cheng and Curtis (2004) identified variables such as the researcher's background, the study's location, and the timing of the assessment, all of which would influence how washback might manifest in practice.

Washback on Teaching

The influence of high-stakes tests on teaching practices has been a central concern in washback research. Much of the research in this area has found that tests tend to affect the content taught, but not always

the methods employed by teachers. Wall and Alderson (1993), for example, found that teachers often adjusted the topics they taught to align with the content of the test, but their instructional methods remained largely unchanged. This suggests that while tests influence "what" is taught, they have a more limited effect on "how" teaching occurs. Early research by Herman and Golan (1991) and Shepard and Dougherty (1991) highlighted how teachers focused on preparing students for tests by using past tests and test preparation materials. Shohamy (1993) examined the impact of Arabic and English language tests, noting that teachers adapted their teaching to match the test content, focusing on vocabulary and oral language. Wall and Alderson (1993) found that despite a revised O-level test in Sri Lanka promoting communicative teaching, teachers narrowed their focus to testable skills like reading and writing.

Studies like Cheng (2008), Watanabe (2004), and Pedulla et al. (2003) further emphasized how external pressure from tests led teachers to concentrate on test-specific content, even if it conflicted with the broader goals of reform. In Qi's (2004) research on the National Matriculation English Test in China, she discovered that the focus on grammar in classrooms contradicted the test's goal of assessing language use. Scott (2005, 2007) and Wall (2005) found similar trends, with teachers focusing on test preparation due to its high-stakes nature. Saif (2006) observed a clear link between a spoken language proficiency test and the methodology adopted by teachers. However, she admitted that only one teacher was involved in the study. These studies collectively reveal that the impact of standardized tests is complex, with teachers often adapting their methods to align with the test, sometimes at the expense of broader educational goals.

In the Libyan context, several studies have explored the washback effects of high-stakes tests on teachers' practices. Abulgasem (2024), in a six-page research paper, examined the pedagogical impact of teachers' classroom tests in secondary schools, using questionnaires administered to 27 teachers. His study showed that teachers relied heavily on achievement tests, but the study lacked qualitative data to deepen the analysis. As a result, the research failed to generalize findings across the broader population of Libyan secondary school teachers. Similarly, Ali (2024) studied preparatory school teachers. However, his study shows a significant misalignment between the research title, abstract, objectives, and research questions.

Abdulhamid (2019) explored the relationship between the alignment of components of the Libyan education system and the washback of the Secondary Education Certificate Examination. Using questionnaires and interviews with three teachers, as well as classroom observation, her study indicated negative washback on some teachers, but little to no washback on others.

Ahmed (2018) investigated the influence of the Secondary Education Certificate on English language teachers' practices in southwest Libya, using case study research. His findings revealed negative washback, highlighting the disconnect between the communicative approach prescribed by the curriculum and the grammar-translation method employed by teachers. Interviews with teachers revealed that they had little knowledge of the communicative approach and focused on teaching test-related content, neglecting skills not tested in the test.

Similarly, Onaiba (2014) examined the impact of the Basic Education Certificate English Exam on teachers' instructional practices, finding that teachers adapted their practices to fit the exam format, neglecting important elements that were not included.

These studies emphasize the importance of understanding how teachers and the broader educational context mediate washback effects. Green (2007) emphasized that teachers' professional training and understanding of test design would be crucial in shaping how washback manifests in practice. Teachers who are well-trained and knowledgeable about the test's purpose are more likely to adapt their teaching methods effectively, even if the test content itself remains narrow. Consequently, while washback may influence what is taught, it does not necessarily transform teaching methodologies.

Research Gaps and the Contribution of This Study

Although existing studies have investigated washback in various educational contexts, including Libya, much of the research has predominantly concentrated on the content of high-stakes tests and their direct influence on the teaching materials covered in the classroom. However, there remains a significant gap in understanding how broader contextual factors—such as teacher professionalism, institutional constraints, and societal influences—mediate the nature and extent of washback. Furthermore, there is a lack of up-to-date research on washback within the Libyan context, particularly concerning the changes introduced by the new curriculum in the late 1990s and the evolving effects of high-stakes testing since then.

This study seeks to address this gap by investigating the influence of high-stakes tests on teaching practices in Libyan preparatory schools. Specifically, it aims to explore how teachers adapt their instructional practices in response to the content of the final test. The study will be guided by Alderson and Wall's (1993) washback hypothesis, which

provides a framework for understanding how tests influence both the "what" and the "how" of teaching. Given that the final test assesses isolated vocabulary and grammatical knowledge, this framework will help examine the potential misalignments between the communicative approach prescribed in the curriculum and the teaching practices shaped by the test. In addition to focusing on the test itself, this study will consider the broader educational context, including institutional and societal factors that may mediate the effects of washback. By doing so, it aims to provide a more comprehensive understanding of washback in the Libyan educational system, offering insights into the contextual variables that shape its impact on teaching practices.

Research Methodology

This study follows the case study approach, which is an empirical inquiry that investigates a contemporary phenomenon within its real-life context using multiple sources of evidence (Yin, 2003). The case study approach allows for a deeper understanding of the washback phenomenon within the context of Libyan preparatory schools. The study employed a qualitative research design within the interpretative paradigm. It allows for an understanding and interpretation of teachers' behaviours and classroom practices influenced by the final test (Bryman, 2008; Cohen et al., 2007; Dornyei, 2007). This design is descriptive (Merriam, 1988; Yin, 2003) and focuses on providing a detailed account of the teachers' perspectives regarding the washback effects of the final test on their teaching practices.

Data collection

The most appropriate and practical instruments used for collecting data for washback research include questionnaires, interviews and classroom observation (Watanabe, 2004). Video-classroom observation and interviews were used for collecting the data for this study. It was carried out in March and April 2010 in three preparatory schools in Rhebat city in Libya.

Participants

The study involved three teachers: one male and two females, who were teaching in preparatory schools in Rhebat city at the time of data collection. To maintain confidentiality, pseudonyms were used for their identification (e.g. Sh.T, A.T, and G.T).

Sh.T held a Licence in English language from a high institute (after finishing the secondary stage) where he studied for four years. He started teaching in 2004. His professional qualifications were: three courses: one in computing and two courses in English language. These courses were provided for them locally by the Education Secretariat in Rhebat in 2008. The first English language course was for two weeks.

It was given by a Libyan teacher and covered linguistics. The second course was on phonetics, and it lasted for three weeks.
A.T had a Licence in English language form a teacher training college (after finishing the secondary school stage). She spent four years in order to obtain her academic qualifications. She began teaching in 2007 (three years). She took one course to achieve a professional qualification. This lasted for a couple of weeks and was given by the Education Secretariat in Rhebat. It was about computing and psychology.
G.T had a specialist teaching licence in English language teaching which she obtained from a middle institute (just after finishing the preparatory stage) where she studied for five years. She taught English for 18 years solely in preparatory schools. Her professional course experience consisted of one course taken for two weeks. It concerned teaching the new course materials.

Sampling of the Schools

A non-probability convenience sampling method was employed to select the schools for this study. This approach was chosen due to ease of access (Bryman, 2008; Cohen et al., 2007; Dornyei, 2007), and the fact that the teachers shared common characteristics relevant to the research aims. The selection of three schools was based on the following criteria:
1. Permission was granted to conduct video-recordings of classroom lessons.
2. The teachers all came from a similar educational and social background, ensuring consistency in the sample.

Interviews

Qualitative interviews were used to gain insights into the teachers' motivations behind their classroom practices. The interviews followed a semi-structured format, allowing for flexibility in the order of questions while covering predefined themes (Mann, 2011; Gray, 2004; Denscombe, 1998). To reduce language barriers and facilitate a more detailed discussion, the interviews were conducted in Arabic.
Two rounds of interviews were conducted in March and April of 2010, and each interview was audio-recorded. Before the interview, the purpose and structure were explained to the participants, and they were informed of the estimated duration of the interview. Afterward, they were given the opportunity to listen to the recordings and provide any additional comments. The interviews focused on two main areas: current teaching practices: Teachers' use of materials, teaching objectives, and the perceived effects of the final test on their teaching methods.
Thematic analysis was employed to analyse the interview data. According to Braun and Clarke (2006), a thematic approach is used for "identifying, analysing and reporting patterns (themes) within data"

(p.79). A theme captures any important issue about the data in relation to the research question(s). The data were reviewed multiple times, and initial codes were generated manually. The codes were then grouped into potential themes, and the relationships among the themes were reviewed. Finally, the themes were refined to ensure coherence and relevance to the research questions.

Classroom Observation

Classroom observations were conducted using a participant observation approach (Bryman, 2008; Walsh, 2006; Denscombe, 1998). This qualitative method allowed the researcher to observe teachers' interactions with students and the implementation of teaching practices. The primary aim of the classroom observation was to assess how teachers applied the principles of the communicative approach and to identify any differences in the content and the methodology of their lessons.
Video-recordings were used to capture the classroom interactions. These recordings were later transcribed verbatim to provide an accurate representation of the lessons. The analysis of these transcripts involved both inductive and deductive coding (Payne & Payne, 2004). Initially, an inductive approach was used to identify patterns and insights related to the washback effects of the test. The analysis then transitioned to a deductive approach to verify whether the principles of the communicative approach were being followed in the observed lessons. The findings from the classroom observations were used to triangulate the interview data and provide a comprehensive understanding of how the final test influenced teachers' practices in the classroom.

Findings

The findings from the interviews and classroom observations are classified into three main themes: 'Teaching Arrangements', 'Objectives of Teaching', and 'Potential Effects of the Final Test on Teaching'.

Teaching Arrangements

The teachers were provided with specific teaching instructions in the teacher's book, which were intended to guide their lessons. However, when asked about their adherence to these instructions, the teachers expressed that they did not fully implement them in their teaching. The main reasons for this were time constraints and the perceived mismatch between the prescribed objectives and the students' abilities. For example, A.T explained that the prescribed content often exceeds her students' comprehension level, leading her to modify or skip some lessons:

A.T: "I don't follow them that much."

I: "Why?"

A.T: "Because they are hardly ever applicable to the students. For example, 'Alhambra Palace' is difficult. It takes me two classes to read it, and by then, the students have only just read it, so when do we answer the questions?"

Despite this, the teachers were required by the education authorities to outline their lesson plans, but they admitted that they did so mainly to satisfy inspectors:

Sh.T: "No, I just write them to satisfy the inspector."

Furthermore, the teachers indicated that they were unfamiliar with the communicative approach to teaching, which might explain why they did not incorporate it into their practices:

I: "Are you familiar with the communicative approach?"

Sh.T: "No, by Allah. What is communicative approach in Arabic?"

The other teacher:

I: "Ok, the course book is designed on the principles of the Communicative Approach do you have any idea about it?"

A.T: "No, no."

Objectives of Teaching

The teachers' stated teaching objectives were often vague and did not consistently reflect their classroom practices. Sh.T mentioned that his primary objective was to prepare students for their future specialties, although this goal was not clearly evident in the observed lessons:

I: "Do you have any particular objectives?"

Sh.T: "Yes, to prepare them for the specialty they will pursue in the future."

A.T described her objective as helping students master English, particularly in pronunciation, as she felt that proper pronunciation was key to language mastery, regardless of grammar rules:

A.T: "The objective is for them to master how to use it. When they pronounce it correctly, it means you don't have to worry about them... even if I don't teach the rules, someone else might."

G.T had a more general view, stating that her objective was for students to acquire reading, writing, and grammar skills:

G.T: "The objective is that they know how to read and write... and grammar."

When asked about their focus during lessons, all three teachers stressed the importance of grammar, reflecting its central role in their teaching and its alignment with the content of the final test. Sh.T explicitly stated that grammar was essential for mastering the English language:

Sh.T: "Yes, I focus on grammar."

I: " Why?"
Sh.T: "Because it is the only exit to the language."

A.T also emphasized the importance of grammar but did not offer a detailed explanation as to why:

A.T: "It's almost the most important in teaching materials."
I: "Why?"
A.T: "I don't know, but umm we apply it."

Potential Effects of the Final Test on Teaching

The teachers' teaching practices appeared to be influenced by the content and format of the final test. A common theme across all three teachers was the exclusion of certain content, such as listening and speaking skills, which were not tested. For example, Sh.T stated that he avoided teaching listening because it would not be included in the final test:

Sh.T: "Listening is the most that I leave out."
I: "Why do you leave it out?"
Sh.T: "It is not included in the final test."

A.T echoed this sentiment, explaining that she left out topics which would not appear in the test to save time:

A.T: "We cover things like 'Dubai' or 'Sea Hotel,' but I know these aren't in the final test."

Similarly, G.T admitted to excluding listening and, at times, writing instruction for the same reason, though she acknowledged that she would teach these skills if they were tested:

G.T: "Listening, sometimes writing, I don't teach them because they are not tested in the final examination."
I: "If they were tested in the final examination?"
G.T: "Then I might teach them."

Furthermore, the teachers' own test design practices were influenced by the final test format. Sh.T, for example, reported that he had designed his tests to mirror those of the final test by using similar questions:

Sh.T: "I design my tests similar to the final test. I want them to get used to the test questions because the same questions are repeated in the tests."

A.T also adapted her tests to resemble the final test format, particularly through multiple-choice questions (MCQs), but adjusted the content to better suit her students' lower proficiency level:

> A.T: "I design MCQs because that's similar to the final test. Sometimes I make changes to my tests to help my students, given their low level."

In contrast, G.T's tests did not directly mirror the final test format. She preferred using questions directly from the course materials, but still excluded listening and speaking from the assessments:

> G.T: "My tests are based on the content of the lessons, not like the final test."
> I: "Do you test writing or speaking?"
> G.T: "No, no."

However, when reviewing her tests, it became apparent that they still shared common features with the final test, such as multiple-choice questions, spelling, and decontextualized vocabulary and grammar.

Classroom Observation

While the course materials were designed according to the principles of the communicative approach, the teachers' classroom methodology predominantly followed the grammar-translation method. They focused on teaching discrete elements of grammar and vocabulary, and predominantly used Arabic in their instruction. In some lessons, the teachers made explicit references to the final test. These references were observed in three lessons taught by A.T, one lesson by G.T, and one lesson by Sh.T.

In March 2010, A.T made a direct reference to the final test when explaining the grammatical rule for changing nouns ending in 'y' into their plural form. She emphasized that this rule frequently appeared in the final test, urging the students to pay attention to it. Later in April, A.T made two additional explicit references to the test. The first occured when she highlights a specific lesson and explains how the final test would assess the content she is teaching. The second reference happened when she expressed frustration that her students did not know the meaning of certain words. She warned them that if they did not learn those words, they would not be able to answer related questions in the final test.

On April 25, 2010, G.T's students initiated a reference to the final test during a classroom activity. While the class was in session, some students began to take out previous copies of the test. One student asked G.T whether they can answer the test questions:

> Student: ") "يا أبله, انحلوا الامتحان؟O teacher, can we answer the test?

G.T: "شنوا الحاجة اللى مش واضحة؟ اسألو What is unclear? Ask.

A.T responded similarly during one of her lessons. She told her students that a particular question always appeared in the test, reinforcing its importance:

A.T: ماهو فى الامتحان من بحله؟ كان مش فاهمين ومعرفتش الاجابة لاكن المهم فهمت" Because it is in the test. Who will answer it? If you don't understand and don't know the answer, what matters is that you understood.

Student: " شن هيا carrot?) "؟What's carrot?"

Although Sh.T did not make an explicit reference to the test in his lessons, he used previous final test questions during classroom revision sessions held in April 2010. This indirect reference to the final test allowed the students to become familiar with the types of questions they might encounter.

The classroom observations highlighted an emphasis on the final test in the teachers' lesson planning and instruction. Despite the communicative approach being part of the curriculum, the teachers primarily adopted a grammar-translation method. They focused on grammar rules and vocabulary, often using Arabic, and explicitly made references to the final test to ensure students were prepared for the test. This suggests that the final test had a significant effect on the teachers' content.

Discussion

Before diving into the discussion, it is important to address two potential criticisms of this study. First, while the data was collected several years ago, this does not invalidate or undermine its reliability, especially since the structure of the test has remained unchanged. Although there have been minor adjustments to the course materials, and the test format (as of 2024) consists of 56 questions—12 True/False and 44 multiple-choice questions, these changes may not affect the study's results. Therefore, the findings might still be considered valid. Second, while the sample size may be small, I am confident that this does not undermine the study's relevance. The data collected was used as a benchmark for future research, and the qualitative nature of the study allowed for a smaller sample. While larger studies could provide more data, it is unlikely that the findings would differ significantly unless we find the occasional outstanding teacher.

The analysis of interviews and classroom observations over three months revealed that the teachers did not fully understand the communicative approach embodied in the course materials, nor did they grasp the purpose of the final test. Classroom activities predominantly reflected a grammar-translation method, with teachers over-relying on Arabic and teacher-fronted instruction. In the interviews, all three teachers admitted that they struggled to teach the course materials communicatively, as they lack understanding of the approach or its application. The teachers' main effort in covering all units was driven by the structure of the final test, which included questions from various course units. While these efforts might seem like a response to the test, they cannot be regarded as positive washback, as they do not support the educational goals (see Hughes, 2003; Bailey, 1996).

Despite the absence of formal accountability mechanisms in Libyan schools (e.g. no direct consequences for funding or rankings), teachers still felt pressure due to internal accountability driven by ethical considerations. This was particularly evident when A.T stated, "you start feeling sympathetic to them. There is justice between us," and when Sh.T referred to the pressure he feels "from myself" to complete the course material. Additionally, teachers faced social pressures related to parents' expectations, as success would often be measured by students' pass rates, as highlighted by A. T's comment:

> A.T: "Yes, it is supposed to find out the students' level, more than pass or fail."
> I: "Yes, yes."
> A.T: "But they look for those who are successful only."

This illustrates how societal norms and expectations influence teachers' behaviours.

Throughout the study period, teachers made explicit references to the final test by directing students' attention to grammar and vocabulary likely to appear on the test, and by using previous test papers. Spolsky (1981) warned that such practices could narrow the curriculum and encourage excessive test preparation. While I agree that these practices have negative implications, I disagree with the conclusions of Abualgasem (2024), Abdulhamid (2019), Ahmed (2018), and Onaiba (2014), who focused on test washback in the Libyan context. This is because the final test predominantly assesses multiple-choice questions on decontextualized grammar and vocabulary, and the teachers are unfamiliar with the communicative approach or its focus. Moreover, there is no evidence at all that these practices are a direct result of the current test design and do not pre-date the test being examined in this paper. All this suggests that cultural factors, such as the teacher's authority and emphasis on memorization, play a strong role in shaping

classroom practices. This aligns with findings from other washback research, which emphasizes the role of the teacher as a mediating factor (Cheng et al., 2015; Green, 2007; Watanabe, 1996, 2004).

Conclusion and Recommendations

The teachers' practices observed in this study are likely representative of many Libyan classrooms. While the influence of high-stakes tests on teachers is widely acknowledged, the extent of this influence is still debated. This study shows that the teachers' lack of knowledge about the communicative approach, combined with the structure of the final test, leads them to focus on grammar and decontextualized vocabulary. These practices cannot be solely attributed to the washback effect of the test, but rather to a lack of understanding of modern language teaching methodologies.

This study extends our understanding of washback research by suggesting that when the test constructs and teachers' practices align, we cannot definitively attribute teachers' behaviours to the washback effect, whether positively or negatively. In this case, the teachers' focus on grammar and vocabulary is not a response to the test itself, but a reflection of their limited understanding of how to teach the language.

The findings open several avenues for future research. As this study reveals, washback research often produces conflicting results. To advance the field, a deeper understanding of washback effects, particularly those related to high-stakes tests, is necessary. Future research could investigate the underlying perceptions and attitudes that drive teachers' actions, particularly examining why discrepancies exist between what teachers say and what they actually do. Additionally, this study could serve as a foundation for exploring the relationship between teachers' understanding of language learning principles, language teaching methodologies, and test design in shaping washback. Given that this study was confined to three rural schools, replicating it in urban schools would help verify or challenge these findings. Moreover, including perspectives from other stakeholders, such as school administrators, inspectors, and test designers, would be valuable in further understanding the washback phenomenon and its broader implication.

A useful follow-up study would involve a comparison of current classroom practices with those identified in this study, including the reasons for any changes (or for no changes).

REFERENCES

Abdulhamid, N. (2019). What is the relationship between alignment and washback? A mixed-methods study of the Libyan EFL context (Doctoral dissertation). *Carleton University*.

Abulgasem, S. (2024). Exploring the pedagogical impact of testing in Libya: Consequences and assessment strategies in English language education. *International Journal of Research Publication and Reviews, 5*(9),1633–1639.

Aldabbus, S. (2008). An investigation into the impact of language games on classroom interaction and pupil learning in Libyan EFL primary classrooms (PhD thesis). *University of Newcastle upon Tyne*.

Alderson, J. C. (2004). Foreword. In L. Cheng, Y. Watanabe, & A. Curtis (Eds.), *Washback in language testing: Research contexts and methods* (p. x). Lawrence Erlbaum Associates.

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

Alderson, J. C., & Hamp-Lyons, L. (1996). TOEFL preparation courses: A study of washback. *Language Testing, 13*(3), 280–297.

Alhmali, R. (2007). Students' attitudes in the context of the curriculum in Libyan education in middle and high schools (PhD thesis). *University of Glasgow*.

Ali, R. E. (2024). Washback in English education: Insights from Libyan ninth graders. *Journal of English Studies in Arabia Felix, 3*(2), 35–43. https://doi.org/10.56540/jesaf.v3i2.94

Bachman, L., & Palmer, A. (1996). *Language testing in practice*. Cambridge University Press.

Bailey, K. M. (1996). Working for washback: A review of the washback concept in language testing. *Language Testing, 13*(3), 257–279.

Barton, L. C. (1968). *Libya: English teaching, 27 August 1965 to 30 June 1968*. Paris: UNESCO (Serial No. 899/BMS.RD/EDM).

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*, 77–101.

Bryman, A. (2008). *Social research methods*. Oxford University Press.

Cheng, L. (2005). Changing language teaching through language testing: A washback study. *Cambridge University Press*.

Cheng, L., & Curtis, A. (2004). Washback or backwash: A review of the impact of testing on teaching and learning. In L. Cheng, Y. Watanabe, & A. Curtis (Eds.), *Washback in language testing: Research contexts and methods* (pp. 3–17). Lawrence Erlbaum Associates.

Cheng, L. (2008). Washback, impact, and consequences. In E. Shohamy & N. H. Hornberger (Eds.), *Encyclopaedia of language and education* (2nd ed., Vol. 7, pp. 349–364). Springer.

Cohen, L., Manion, L., & Morrison, K. (2001). *Research methods in education*. Routledge.

Cumming, A. (2009). Language assessment in education: Tests, curricula, and teaching. *Annual Review of Applied Linguistics, 29*, 90–100.

David, M., & Sutton, C. D. (2004). *Social research: The basics*. SAGE.

Denscombe, M. (1998). *The good research guide for small-scale social research projects*. Open University Press.

Dцrnyei, Z. (2007). *Research methods in applied linguistics: Quantitative, qualitative, and mixed methodologies*. Oxford University Press.

Fredericksen, N. (1984). The real test bias: Influences on testing and teaching. *American Psychologist, 39*, 193–202.

Gray, D. E. (2004). *Doing research in the real world*. SAGE.

Green, A. (2007). Washback to learning outcomes: A comparative study of IELTS preparation and university pre-sessional language courses. *Assessment in Education, 3*(4), 333–367.

Hughes, A. (2003). *Testing for language teachers*. Cambridge University Press.

Linn, R. L. (2000). Assessment and accountability. *Educational Researcher, 29*(2), 4–16.

Madaus, G. F. (1988). The influence of testing on the curriculum. In L. Tanner (Ed.), *Critical issues in curriculum* (pp. 83–121). Chicago University Press.

Mann, S. (2011). Critical review of qualitative interviews in applied linguistics. *Applied Linguistics, 3*(1), 6–24.

Merriam, S. (1988). *Case study research in education*. Jossey-Bass.

Messick, S. (1996). Validity and washback in language testing. *Language Testing, 13*(3), 241–256.

Onaiba, A. (2014). Investigating the washback effect of a revised EFL public examination on teachers' instructional practices, materials, and curriculum (Doctoral dissertation). *University of Leicester*.

Orafi, S., & Borg, S. (2009). Intentions and realities in implementing communicative curriculum reform. *System, 37*, 243–253.

Payne, G., & Payne, J. (2004). *Key concepts in social research*. SAGE.

Pearson, I. (1988). Tests as levers for change. In R. Baumgardner & D. Chamberlain (Eds.), *ESP in the classroom: Practice and evaluation* (pp. 98–107). Oxford Modern English Publications in association with the British Council.

Pedulla, J. J., Abrams, L., Madaus, G. F., Russell, M. K., Ramos, M. A., & Miao, J. (2003). Perceived effects of state-mandated testing programs on teaching and learning: Findings from a national survey of teachers. *National Board on Educational Testing and Public Policy*.

Popham, W. (1987). The merits of measurement-driven instruction. *Phi Delta Kappa, 68*, 679–682.

Qi, L. (2004). Has a high-stakes test produced the intended changes? In L. Cheng, Y. Watanabe, & A. Curtis (Eds.), *Washback in language testing: Research contexts and methods* (pp. 171–190). Lawrence Erlbaum Associates.

Rea-Dickins, P., & Scott, C. (2007). Washback from language tests on teaching, learning, and policy: Evidence from diverse settings. *Assessment in Education: Principles, Policy & Practice, 14*, 1–7.

Saif, S. (2006). Aiming for positive washback: A case study of international teaching assistants. *Language Testing, 23*(1), 1–34.

Saville, N. D. (2009). Developing a model for investigating the impact of language assessment within educational contexts by a public examination provider (PhD thesis). *University of Bedfordshire*.

Scott, C. (2005). Washback in the UK primary context with EAL learners: Exploratory case studies (Unpublished doctoral dissertation). *University of Bristol*.

Shihiba, S. (2011). An investigation of Libyan EFL teachers' conceptions of the communicative learner-centred approach in relation to their implementation of an English language curriculum innovation in secondary schools (PhD thesis). *Durham University*.

Shohamy, E. (2001). *The power of tests: A critical perspective on the uses of language tests*. Longman.

Spolsky, B. (1981). Some ethical questions about language testing. In C. Klein-Braley & D. K. Stevenson (Eds.), *Practice and problems in language testing* (pp. 5–21). Peter Lang.

Spratt, M. (2005). Washback and the classroom: The implications for teaching and learning of studies of washback from examinations. *Language Teaching Research, 9*(1), 5–29.

Taylor, L. (2005). Washback and impact. *ELT Journal, 59*(2), 154–155.

Wall, D. (2005). The impact of high-stakes examinations on classroom teaching: A case study using insights from testing and innovation theory. *Cambridge University Press*.

Wall, D., & Alderson, C. (1993). Examining washback: The Sri Lankan impact study. *Language Testing, 10*(1),Wallace, M. J. (1998). *Action research for language teachers*. Cambridge University Press.

Walsh, S. (2006). *Investigating classroom discourse*. Routledge.

Watanabe, Y. (2004). Teacher factors mediating washback. In L. Cheng, Y. Watanabe, & A. Curtis (Eds.), *Washback in language testing: Research contexts and methods* (pp. 129–146). Lawrence Erlbaum Associates.

Yin, R. (2003). *Case study research: Design and methods* (3rd ed.). SAGE.