



Why Teachers Do Not Teach for Understanding: The Washback Effects of SSC Exams on Pedagogical Practices of Teachers in Sargodha

Abdul Razzaq¹, Sadaf Kiran², Muhammad Usman Sajjad³, Muhammad Ali Shahid⁴

¹ Lecturer, Government Graduate College Sillanwali, Sargodha. Email: kashanaakash@gmail.com

² Lecturer Government Associate College for Women, Choubara, Layyah.

Email: sadafkiran3@gmail.com

³ SSE, Government High School 113 NB, Sillanwali. Email: usmansajjad550@gmail.com

⁴ Lecturer, department of English Language and Literature the University of Lahore Sargodha, Campus. muhammadalishahid05@gmail.com

DOI: <https://doi.org/10.70670/sra.v4i1.1858>

Abstract

High-stakes examinations exert a powerful influence on classroom instruction, particularly in developing-country contexts where public assessments dominate educational culture. This qualitative case study investigates the washback effect of the Secondary School Certificate (SSC) examinations administered by the Board of Intermediate and Secondary Education (BISE), Sargodha, on the pedagogical practices of secondary school teachers. Data were gathered through semi-structured interviews with five teachers and six students from a government high school in District Sargodha and through the systematic analysis of five years of SSC examination papers using Bloom's Taxonomy revised framework. Findings reveal that between 88% and 92% of examination questions target only the knowledge (recall) level of Bloom's Taxonomy, with negligible representation of higher-order thinking skills (HOTS) such as analysis, synthesis, or evaluation. Consequently, teachers adopt examination-oriented pedagogical methods centred on rote memorisation, repeated practice of past papers, and guidebook dependency, all of which foreclose meaningful language development. Students similarly prioritise memorised answers over communicative competence. The study concludes that the SSC examination generates strong negative washback and recommends systemic assessment reform encompassing formative assessment, performance-based tasks, and higher-order thinking questions to restore pedagogical quality and promote positive washback.

Keywords: Washback Effect, SSC Examination, Assessment Practices, Rote Memorisation, Higher-Order Thinking, Pedagogy, Pakistan

Introduction

High-stakes testing occupies a uniquely powerful position in shaping educational practice worldwide. When public examinations function as gatekeepers to higher education, they inevitably reorient classroom instruction toward test preparation, often at the expense of substantive intellectual development. This alignment of pedagogy with assessment demands—commonly described as "teaching to the test"—represents one of the most widely documented phenomena in educational research (Bailey, 1996). In Pakistan, the Secondary School Certificate (SSC) examination, administered by regional Boards of Intermediate and Secondary Education (BISEs) at the end of Grade 10, constitutes precisely such a high-stakes threshold. The examination determines access to intermediate-level education and carries enormous social and institutional

weight, making it a decisive influence on everything teachers do inside the classroom (Mahmood, 2012). The consequences are well documented in the broader literature but remain underexplored in the specific context of BISE Sargodha. Teachers who work under the shadow of a demanding public examination frequently abandon learner-centred, inquiry-based, or communicative approaches in favour of methods that reliably produce examination success. Rote memorisation of textbook passages, stock answers, and previously administered questions becomes the dominant learning strategy (Shamim, 2011). Rather than cultivating the analytical reasoning, critical evaluation, and creative synthesis that contemporary educational philosophy emphasises—skills comprehensively articulated in the revised framework of Bloom's Taxonomy (Anderson & Krathwohl, 2001)—classrooms are transformed into rehearsal spaces for a single, high-pressure performance.

This situation is particularly concerning in the context of English language education. The SSC English paper is intended to develop communicative language competence across the four macro-skills of reading, writing, listening, and speaking. Yet the examination format, as will be demonstrated in this study, overwhelmingly privileges reading comprehension recall and written reproduction of memorised material. Listening and speaking are virtually absent from formal assessment. When assessment defines learning, and when that assessment measures only a narrow slice of language competence, the long-term consequences for students' functional English proficiency are severe and frequently irreversible by the time students reach tertiary education.

The present study was therefore designed to examine the washback effect of SSC examinations on classroom pedagogy in a government secondary school in District Sargodha. It pursues three interrelated objectives: to document how examination requirements shape teachers' instructional decisions; to determine which cognitive levels SSC examination questions actually test; and to capture the perceptions of both teachers and students concerning examination-oriented learning. In pursuing these objectives, the study aims to contribute empirical evidence from an underrepresented Pakistani context to the global washback literature and to generate practically grounded recommendations for assessment reform.

Literature Review

The Concept of Washback

Washback—also referred to as backwash in some British scholarship—denotes the influence that tests exercise on teaching and learning. The concept entered applied linguistics as a formal object of inquiry in the early 1990s, most influentially through the work of Alderson and Wall (1993), whose landmark paper questioned whether tests inevitably influence teaching and, if so, under what conditions and through what mechanisms. Their "Washback Hypothesis" proposed a series of nuanced claims: that tests may influence what and how teachers teach, what and how learners learn, the rate and sequence of teaching and learning, the degree and depth of teaching and learning, attitudes toward content and methods, and the nature of preparation for tests. Crucially, Alderson and Wall (1993) resisted a simplistic causal account, arguing instead that the relationship between testing and pedagogy is mediated by a complex web of contextual, institutional, and individual variables.

Bachman and Palmer (1996) subsequently offered a broader conception of washback as encompassing the effects of tests on educational systems beyond the classroom, including curriculum design, materials development, and institutional policy. They characterised washback as "a complex phenomenon" that extends well beyond straightforward test-teaching relationships. Bailey (1996) synthesised the emerging literature and identified four principal mediating variables: participants (teachers and learners), processes (instructional and learning activities), products (learning outcomes and competencies), and test characteristics (format, content, and stakes). This multi-variable model remains influential because it acknowledges that identical tests may produce very different washback effects depending on who takes them and in what institutional setting.

More recent bibliometric analysis confirms that washback research has expanded substantially over the past

three decades, with a marked shift toward English language testing contexts in Asia and the Middle East (Xie & Jia, 2025). Ahmad et al. (2025) examined the washback of the Higher Secondary School Certificate English Language Examination (HSSC-ELE) in Kohat District and found a pattern strikingly consistent with the present study: teachers systematically structure lessons around anticipated examination questions rather than broader communicative or curricular goals. Cheng (2005) demonstrated through longitudinal research in Hong Kong that changes to examination design produce corresponding, though often lagged, changes in teaching practice, confirming that assessment reform can be a lever for pedagogical improvement if implemented carefully.

Positive and Negative Washback

The washback literature distinguishes consistently between positive and negative effects. Hughes (1989) framed this distinction in terms of whether a test promotes or undermines effective learning. Positive washback occurs when examinations are designed to assess authentic language use, higher-order cognitive skills, or communicative performance, thereby incentivising instruction that develops those capacities. Negative washback occurs when examinations assess only surface-level knowledge through formats—such as decontextualised recall questions and closed multiple-choice items—that reward memorisation and penalise the time investment required for deeper learning.

Green (2007) added analytical precision to this distinction by noting that the direction of washback is substantially determined by examination design. Assessments that target communicative competence and higher-order thinking are more likely to generate positive washback, while those that operationalise achievement through recall metrics reliably produce negative effects. Christiana (2025) corroborated this in a recent literature review, arguing that the nature of examination washback depends fundamentally on what the assessment values and how it operationalises proficiency or learning.

A related distinction distinguishes strong from weak washback (Cheng et al., 2004). Strong washback denotes a direct and powerful influence of examinations on teaching behaviour; weak washback implies that assessments exert only marginal pressure, allowing teachers latitude to pursue comprehensive pedagogical goals. The strength of washback typically correlates with the stakes attached to the examination: the higher the institutional and personal consequences of performance, the stronger the washback (Green, 2007). In the SSC context—where the examination determines academic trajectory and carries significant social prestige—strong washback is the predictable outcome.

Washback can also be characterised as overt or covert. Overt washback is visible in deliberate examination preparation activities—past paper drilling, guidebook memorisation, topic prediction—while covert washback operates through subtler shifts in teacher beliefs, professional identity, and values about what education is for. Dammak et al. (2023) documented both dimensions in their study of English Baccalaureate washback in Tunisia, finding that prolonged exposure to high-stakes examination culture alters not merely what teachers do but how they conceive of their professional role.

The SSC Examination System in Punjab

The Secondary School Certificate examination in Punjab is administered by eight regional Boards of Intermediate and Secondary Education. Candidates sit for two papers: Paper I at the end of Grade 9 and Paper II at the end of Grade 10. Each paper comprises an objective section (multiple-choice questions, constituting 20% of the total marks), a short-answer section (50%), and a long-answer section (30%). Questions are theoretically distributed across the cognitive domains of knowledge, understanding, and application, in accordance with Bloom's Taxonomy. In practice, however, as the present study demonstrates, the application and higher cognitive domains receive negligible representation.

The examination is syllabus-bound, meaning that all questions are drawn from designated textbooks within a prescribed syllabus. This boundedness has significant implications for pedagogy: teachers who align

instruction with the textbook are, in effect, covering the entire domain of assessable content, leaving no institutional incentive to extend beyond it. The format privilege of the written examination further constrains what is assessed, effectively marginalising speaking and listening skills from the assessment landscape and, by extension, from classroom attention.

Khan (2016) documented the broader systemic consequences of the SSC examination on teaching and learning at secondary level in Pakistan, finding that teachers across subjects reported modifying their instructional choices directly in response to examination requirements. Mahmood (2012) situates these findings within a historical analysis of Pakistani examination culture, noting that the examination system has evolved in ways that reinforce rather than challenge rote learning tendencies.

Assessment Reform and Washback Mitigation

The global literature on examination reform suggests that washback can be reoriented from negative to positive through deliberate changes in assessment design. Black and Wiliam (1998) demonstrated in their foundational review that formative assessment—ongoing, feedback-rich evaluation integrated into classroom instruction—substantially improves student learning outcomes and shifts educational cultures away from summative anxiety. Harlen and James (1997) argue that genuine learning, defined as the capacity to apply knowledge in novel contexts, occurs when assessment is oriented toward understanding rather than recall. Gipps (1996) proposed performance-based assessment as a constructivist alternative to traditional examinations, arguing that measuring learners' ability to solve problems and accomplish real-world tasks produces more valid and educationally beneficial assessment.

In the language testing context, Hughes (1989, 2003) argued for direct and authentic assessment methods that mirror the language tasks learners will encounter outside the classroom, contending that indirect tests—such as discrete-point multiple-choice items—do not reliably predict real language proficiency and generate systematically misleading washback. Raza (2009) noted that the Pakistani examination system's heavy reliance on indirect testing methods contributes to the disconnect between formal qualifications and functional language ability that employers and tertiary institutions regularly observe.

Research Objectives and Questions

The study was guided by three research objectives:

1. To investigate the washback effect of SSC examinations on the pedagogical practices of teachers.
2. To examine the cognitive levels of SSC examination questions utilising Bloom's revised Taxonomy.
3. To explore the perceptions of teachers and students regarding examination-oriented learning.

These objectives generated three corresponding research questions:

1. What is the effect of SSC examinations on the pedagogical practices of teachers?
2. Which cognitive levels are predominantly tested in SSC examination papers?
3. What are the perceptions of teachers and students regarding the role of examinations in classroom learning?

Methodology

Research Design

A qualitative case study design was employed to investigate the washback effect of the SSC examination on pedagogical practice. Qualitative methodology was selected because the study sought to understand the lived experiences, perceptions, and contextually embedded decisions of teachers and students—phenomena that resist quantification and require interpretive, emic engagement. Case study design, as articulated by Yin (2018), is appropriate when the research aims to understand a contemporary phenomenon within its real-world context and when the boundary between phenomenon and context is not clearly self-evident. The SSC

examination system and the single school under investigation together constitute a bounded case that satisfies these criteria.

The study was conducted at a Government Boys High School in District Sargodha, Punjab, Pakistan. The site was selected through purposive sampling on the grounds that it is representative of mainstream public secondary education in the district and that the researcher had established access relationships that facilitated candid participant engagement.

Participants

The participant sample comprised five teachers and six students. Teacher participants were selected purposively to reflect variation in subject area and teaching experience, ensuring that the data captured a range of perspectives on examination-driven pedagogy rather than a single disciplinary or experiential viewpoint. Student participants were selected from Grade 10 to ensure direct relevance to SSC examination preparation. In total, eleven participants were interviewed. Although the sample is necessarily small—a recognised limitation of qualitative case study research—it is consistent with the interpretive aims of the study, which prioritise depth of understanding over statistical generalisation.

Data Collection

Two data sources were employed. First, semi-structured individual interviews were conducted with all eleven participants. Semi-structured interviewing was chosen because it combines the systematic coverage of predetermined themes with the flexibility to pursue emergent lines of inquiry through follow-up probes. Interview questions addressed participants' experiences of examination preparation, their perceptions of the examination's cognitive demands, and their views on the relationship between assessment and learning. Interviews were conducted in participants' preferred language (Urdu or English), recorded with consent, and transcribed for analysis.

Second, SSC examination papers from 2019 to 2023—five consecutive years—were collected and subjected to systematic content analysis using Bloom's revised Taxonomy (Anderson & Krathwohl, 2001) as the analytical framework. Each examination question was independently classified into one of six cognitive levels: Remember, Understand, Apply, Analyse, Evaluate, or Create. The classification was conducted by the researcher and verified through inter-rater review to ensure coding reliability. Results were aggregated to generate percentage distributions of cognitive levels across each examination year.

Data Analysis

Interview data were analysed thematically, following the six-phase process described by Braun and Clarke (2006): familiarisation with the data, generation of initial codes, clustering of codes into candidate themes, review of themes against the data set, definition and naming of final themes, and production of the written analysis. Analysis was inductive in its initial phases and deductive in its later stages, with the washback framework of Bailey (1996) serving as an organising structure for interpreting emergent themes. Member checking—sharing preliminary interpretations with participants—was employed as a validity strategy.

Findings

Cognitive Level Analysis of SSC Examination Papers

The systematic analysis of SSC examination papers from 2019 to 2023 produced consistent and striking results. As shown in Table 1, the overwhelming majority of examination questions in every year target the knowledge (recall) level of Bloom's revised Taxonomy. The proportion of recall-level questions ranged from 88% to 92%, with understanding-level questions accounting for the remaining 8% to 12%. Higher-order cognitive skills—application, analysis, evaluation, and creation—were entirely absent across all five years of examination papers.

Table 1*Cognitive Level Distribution of SSC Examination Questions (2019–2023)*

Year	Knowledge (%)	Understanding (%)	Higher-Order Skills (%)
2019	91	9	0
2020	88	12	0
2021	90	10	0
2022	92	8	0
2023	89	11	0

Note. Higher-order skills encompass the Apply, Analyse, Evaluate, and Create levels of Bloom's revised Taxonomy (Anderson & Krathwohl, 2001).

These findings are unambiguous. The SSC examination as administered in Sargodha makes no cognitive demand above the understanding level and directs the vast majority of its assessment energy toward simple factual recall. For students and teachers whose educational ambitions are channelled through this examination, such a cognitive profile sends a powerful signal: remembering is what matters, and higher-order intellectual engagement is not merely unvalued but untested.

Findings from Teacher Interviews

Thematic analysis of teacher interviews produced four principal themes: examination-oriented teaching as the default mode of instruction; pervasive reliance on rote memorisation; systematic neglect of communicative language skills; and dependence on past examination papers as primary teaching materials. These themes are addressed in sequence below.

All five teacher participants reported that examination preparation functions as the organising principle of their classroom instruction. Rather than deriving lesson content from broader curricular goals or language learning objectives, teachers described working backwards from examination requirements—identifying frequently tested topics, predicting likely questions, and allocating instructional time proportionally to examination weight. One teacher stated: "It is our primary task to prepare students for the board examination. By placing too much emphasis on discussion or conceptual learning, students would not perform well in the exam." Another explained: "Due to the fact that the exam is generally composed of questions based on recall, teachers must focus on memorisation." A third reported: "The majority of students request important questions and anticipated exams. They desire content that will enable them to achieve marks as opposed to learning the subject matter."

These statements reveal a dynamic in which the examination, teachers, and students form a mutually reinforcing system. Teachers teach to the test because students demand it; students demand it because they correctly perceive that examination performance, not learning, is the institutional currency. The examination thus functions as both cause and justification for a pedagogy that would, in other contexts, be considered educationally inadequate.

The theme of rote memorisation appeared across all five teacher interviews. Teachers described their instruction as centred on enabling students to reproduce designated answers accurately rather than to develop understanding. Past examination papers featured prominently as teaching tools: teachers reported using previous years' papers not merely for revision but as the primary vehicle for identifying and transmitting

assessable content. Guidebooks—commercially produced study aids that compile expected questions and model answers—were widely recommended to students as essential preparation resources. As one teacher noted, the guidebooks effectively perform the teacher's analytical role, pre-digesting the textbook into an examination-ready format that students can memorise without engaging with the underlying material. Communicative language skills—speaking and listening—received minimal or no instructional attention. All teacher participants acknowledged that the examination does not assess these skills and that instructional time is accordingly not allocated to them. This finding aligns with the assessment-driven logic identified above: if the examination does not test it, teaching it is considered an inefficient use of examination preparation time. The consequence is that students complete two years of secondary English study without substantive development of the oral communicative competencies that tertiary education and professional life will require.

Findings from Student Interviews

Student interview data corroborated and extended the teacher findings. All six students described their examination preparation as primarily memorisation-based, drawing heavily on guidebooks and past papers. One student summarised the dominant approach: "We tend to memorise answers provided in the guidebooks since it is the same question that turns up in the exam." Another student's account revealed an awareness of the tension between understanding and examination success, but resolved it in favour of strategic memorisation: "It is better to understand, but it is not risky to memorise the answers which are expected to score high marks."

Several students expressed reservations about the English language examination as a measure of their actual language ability, reporting that they perceived little improvement in their functional English proficiency over the two years of SSC study. This perception of a disconnect between examination performance and genuine learning represents a significant finding: students are not merely passive victims of an impoverished assessment system; they are aware of its limitations and have made pragmatic peace with them. Learning, in this context, has been redefined as examination passing, and students have internalised that redefinition.

Students also reported experiencing anxiety about diverging from examination-approved content. The guidebook's model answers function as authoritative texts: departing from them, even in the direction of a more sophisticated or original response, is perceived as risky because examiners expect the rehearsed answer. This dynamic effectively punishes students who have developed genuine understanding and rewards those who have invested in memorisation, inverting the relationship between learning and achievement that educational assessment is designed to uphold.

Discussion

The findings of this study provide clear empirical evidence for a strong negative washback effect generated by the SSC examination on pedagogical practice in Sargodha. The discussion below situates these findings within the existing literature and draws out their implications across the principal dimensions of washback identified by Bailey (1996): examination characteristics, teacher behaviour, student learning strategies, and educational products.

Examination Characteristics and Cognitive Narrowness

The Bloom's Taxonomy analysis demonstrates that SSC examination papers systematically assess only the lowest cognitive levels, with 88%–92% of questions targeting recall and no questions at all targeting application or above. This finding resonates with the broader Pakistani testing literature (Mahmood, 2012; Raza, 2009) and confirms what teachers and students already know through experience: success in the SSC examination requires memory, not understanding. The validity of the examination as an assessment of educational achievement is accordingly compromised. Brown (2000) argues that an examination system of this kind generates adverse washback precisely because it measures an impoverished proxy for the educational

goals it purports to assess. The SSC examination measures whether students can remember designated content; it does not measure whether they can use language, solve problems, evaluate evidence, or think critically—yet it is these latter capacities that education is, in principle, designed to develop.

The prevalence of multiple-choice questions in the objective section of the examination exacerbates this problem. Multiple-choice formats, as Davies et al. (1999) observe, typically operationalise recall and recognition rather than productive language use or analytical reasoning. Students who understand that the objective section rewards recognition of a correct alternative have every incentive to practise recognising correct answers rather than constructing them—a habit of mind antithetical to the development of productive linguistic or analytical competence.

Teacher Behaviour and Pedagogical Constriction

The teacher interview findings demonstrate that examination requirements directly and substantially constrict pedagogical choice. Teachers do not simply supplement their instruction with examination preparation; examination preparation has displaced instruction as the primary classroom activity. This constriction operates along several dimensions. First, it narrows content: teachers restrict themselves to textbook-designated topics that are likely to appear on the examination and avoid extending beyond the syllabus. Second, it narrows method: communicative, interactive, and inquiry-based approaches are abandoned in favour of transmission-based delivery of memorisable content. Third, it narrows professional identity: teachers describe their role as examination preparation rather than language education, suggesting that sustained exposure to examination-driven pedagogy has reshaped their professional self-conception, consistent with Dammak et al.'s (2023) findings on examination washback and teacher identity in Tunisia.

Lyons (1998) characterises this as a conflict between pedagogical principles and examination requirements that most teachers, under institutional pressure, resolve in favour of the examination. Bailey (1996) notes that teachers typically exhibit strong washback when they believe examination-based instruction offers the most reliable path to student success, even when they personally recognise its educational inadequacy. The present data confirm this pattern: several teachers expressed awareness that their examination-oriented methods do not develop genuine language competence while simultaneously insisting that they have no practical alternative.

Student Learning Strategies and Passive Learning

Student learning strategies reflect the examination environment with remarkable fidelity. Guidebook dependency, past paper rehearsal, and answer memorisation are not aberrations but rational adaptations to an assessment system that rewards them. Hughes (1993, as cited in Green, 2007) argues that the product of an examination should be measured by whether it promotes effective learning. By this criterion, the SSC examination fails: students report no meaningful improvement in English communicative ability over two years of study, and they have internalised a conception of language learning as the memorisation of examination answers rather than the development of communicative competence.

The finding that students are aware of the tension between understanding and memorisation—and consciously choose memorisation as the safer strategy—is particularly significant. It suggests that the negative washback of the SSC examination operates not merely through habit or unreflective compliance but through rational calculation. Students correctly perceive that the examination rewards memorisation, and they act accordingly. Reforming their learning strategies therefore requires reforming the incentive structure of the examination itself, not merely providing better instruction within the existing examination framework.

Communicative Skills and the Four-Skill Gap

The consistent neglect of speaking and listening skills documented in this study represents a structural failure of the assessment system. The SSC examination's exclusive reliance on written formats means that two of the

four language macro-skills are never assessed and therefore never taught systematically. The implications extend beyond the SSC years: students who proceed to intermediate and tertiary education with underdeveloped oral communicative competence face systematic disadvantage in academic and professional contexts where spoken English is expected. Assessment reform that integrates speaking and listening—through oral components, performance tasks, or integrated skills assessment—would not only address this gap directly but would generate positive washback incentivising instruction in communicative language skills.

Recommendations

The findings of this study generate five principal recommendations for assessment reform at the SSC level in Punjab. These recommendations are grounded in the washback literature and in the specific limitations of the current examination system identified by this study.

The first recommendation concerns the integration of formative assessment into school-level evaluation practice. As Black and Wiliam (1998) demonstrate, formative assessment substantially improves learning outcomes by providing regular, actionable feedback that enables students to understand their progress and adjust their strategies. In the current SSC context, assessment is exclusively summative: students receive a single set of marks at the end of Grade 9 and Grade 10, with no systematic feedback on their developing understanding. Introducing school-based formative assessment—quizzes, drafts, portfolio tasks, and oral presentations evaluated against transparent criteria—would create a parallel assessment culture that values learning over performance and rewards understanding over memorisation.

The second recommendation involves the redesign of SSC examination content to include higher-order thinking questions at the application, analysis, and evaluation levels of Bloom's revised Taxonomy. The current absence of higher-order questions represents not merely a missed opportunity but an active endorsement of shallow learning. Examination boards in jurisdictions that have successfully reformed assessment—including Hong Kong (Cheng, 2005) and parts of the United Arab Emirates—have demonstrated that introducing higher-order questions generates corresponding shifts in classroom instruction toward more cognitively demanding pedagogy. Introducing even a modest proportion—a minimum of 30%—of higher-order questions in SSC papers would send a powerful signal to teachers that surface-level coverage is no longer sufficient for examination success.

The third recommendation pertains to the introduction of performance-based assessment tasks. Gipps (1996) defines performance assessment as measuring students' capacity to apply knowledge and skills to real-world problems. In the English language context, this might take the form of oral presentations, pair or group speaking tasks, extended writing on authentic topics, or reading tasks based on unseen texts. Hughes (2003) argues that direct assessment of this kind not only measures language ability more validly than indirect methods but generates positive washback by incentivising the development of real communicative competence. Piloting performance-based components within school-based assessments as a precursor to their integration into the formal SSC framework would be a pragmatically realistic reform pathway.

The fourth recommendation addresses the balance of language skills assessed by the examination. The SSC English paper should be restructured to assess all four macro-skills—reading, writing, listening, and speaking—in proportions that reflect their relative importance in educated language use. Listening and speaking components could initially be introduced as school-based internal assessments contributing a modest percentage of the final grade, allowing teachers and institutions to develop the necessary assessment infrastructure before full integration into the public examination.

The fifth recommendation involves teacher professional development. Even well-designed assessment reforms will fail to generate positive washback if teachers lack the pedagogical knowledge and skills to capitalise on the new assessment demands. Sustained professional development in communicative language teaching, higher-order questioning, task-based instruction, and formative assessment practice should accompany any examination reform. Teachers who currently organise their instruction around rote

memorisation are not doing so because they are unaware of better alternatives; many are doing so because, within the current incentive structure, it is the rational professional choice. Changing that incentive structure through examination reform, and supporting teachers to respond to the new incentives through professional development, constitutes the most promising pathway to genuine pedagogical improvement.

Limitations and Future Research

Several limitations of the present study should be acknowledged. The sample was small—five teachers and six students from a single school—and the findings cannot be statistically generalised to the broader population of SSC teachers and students in Sargodha or Punjab. The study relied entirely on self-report data for its account of classroom practice; no classroom observations were conducted, which means that the reported practices, though consistent and convergent across participants, cannot be independently verified. Some teachers indicated involvement in creative tasks for untested skills, a claim that could not be assessed without observational data.

The study also focused exclusively on teachers and students, omitting the perspectives of other stakeholders whose decisions shape the examination system: examination setters, curriculum developers, BISE administrators, and educational policymakers. Future research should employ larger and more diverse samples, integrate classroom observation as a primary data source, and engage the full range of stakeholders whose decisions collectively constitute the washback ecology of the SSC system. Longitudinal designs that track teachers and students across multiple examination cycles would provide a richer account of how washback effects develop and intensify over time.

Comparative research examining washback patterns across different BISEs in Punjab, or across different subject areas within the same BISE, would help identify whether the patterns documented here are specific to Sargodha or systemic across the province. Cross-national comparison with washback studies from similar educational contexts in Bangladesh, India, or Nepal could yield insights applicable to assessment reform across South Asia. As Harlen and James (1997) observe, the ultimate measure of any assessment reform must be whether it promotes learning that transfers to new contexts—and achieving this goal in the SSC system will require sustained, evidence-based, system-wide reform effort.

Conclusion

This study has investigated the washback effect of SSC examinations administered by BISE Sargodha on the pedagogical practices of secondary school teachers. The evidence converges on a clear conclusion: the SSC examination exerts a strong negative washback effect on teaching and learning. Examination papers from 2019 to 2023 reveal that between 88% and 92% of questions target only the recall level of Bloom's revised Taxonomy, with no questions assessing higher-order cognitive skills. In direct response to this examination design, teachers have adopted pedagogical practices centred on rote memorisation, past paper drilling, and guidebook dependency, abandoning communicative, inquiry-based, and skills-integrated approaches. Students have internalised the examination's values, consciously prioritising memorised answers over genuine language development.

The implications are serious and extend beyond examination performance. Students who complete secondary education in this system do so with underdeveloped communicative competence, limited capacity for analytical and critical thinking, and a conception of learning as memorisation rather than understanding. Addressing these outcomes requires not incremental adjustment but systemic assessment reform: the introduction of higher-order thinking questions, formative and performance-based assessment, integration of all four language skills, and sustained teacher professional development. The washback literature is unambiguous that examination design drives pedagogical practice. Changing what the examination values and assesses is the most powerful available lever for improving what happens in Pakistani classrooms. The present study provides empirical foundation for that reform agenda.

References

- Ahmad, K., Sadia, S., & Alam, M. (2025). Exam-based pedagogy: Washback of HSSC-ELE in teaching English language in Kohat District, Pakistan. [Journal name pending publication details].
- Alderson, J. C., & Wall, D. (1993). Does washback exist? *Applied Linguistics*, 14(2), 115–129. <https://doi.org/10.1093/applin/14.2.115>
- Anderson, L. W., & Krathwohl, D. R. (Eds.). (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. Longman.
- Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice: Designing and developing useful language tests* (Vol. 1). Oxford University Press.
- Bailey, K. M. (1996). Working for washback: A review of the washback concept in language testing. *Language Testing*, 13(3), 257–279. <https://doi.org/10.1177/026553229601300303>
- Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, 5(1), 7–74. <https://doi.org/10.1080/0969595980050102>
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp063oa>
- Brown, J. D. (2000). University entrance exams: How to have positive washback on teaching English as a language. *JALT Testing & Evaluation SIG Newsletter*, 4(2), 9–14.
- Cheng, L. (1997). How does washback influence teaching? Implications for Hong Kong. *Language and Education*, 11(1), 38–54. <https://doi.org/10.1080/09500789708666717>
- Cheng, L. (2005). *Changing language teaching through language testing: A washback study*. Cambridge University Press.
- Cheng, L., Watanabe, Y., & Curtis, A. (Eds.). (2004). *Washback in language testing: Research contexts and methods*. Lawrence Erlbaum Associates.
- Christiana, M. (2025). Washback in language testing: A literature review of positive and negative influences on teaching and learning. *Jurnal Pendidikan Bahasa*, 15(3), 397–402.
- Dammak, H., Khatibi, A., & Azam, S. (2023). Attitudes and effectiveness beliefs of teachers of English language in Tunisia. *International Journal of English Linguistics*, 13(4), 1–12. <https://doi.org/10.5539/ijel.v13n4p1>
- Davies, A., Brown, A., Elder, C., Hill, K., Lumley, T., & McNamara, T. (1999). *Dictionary of language testing*. Cambridge University Press.
- Fullilove, J. (1992). The tail that wags. *Institute of Language in Education Journal*, 9, 131–147.
- Gipps, C. (1996). Assessment for learning. In A. Little & A. Wolf (Eds.), *Assessment in transition* (pp. 251–262). Pergamon.
- Green, A. (2007). Washback to learning outcomes: A comparative study of IELTS preparation and university pre-sessional language courses. *Assessment in Education: Principles, Policy & Practice*, 14(1), 75–97. <https://doi.org/10.1080/09695940701272880>
- Harlen, W., & James, M. (1997). Assessment and learning: Differences and relationships between formative and summative assessment. *Assessment in Education: Principles, Policy & Practice*, 4(3), 365–379. <https://doi.org/10.1080/0969594970040304>
- Hughes, A. (1989). *Testing for language teachers*. Cambridge University Press.
- Hughes, A. (2003). *Testing for language teachers* (2nd ed.). Cambridge University Press.
- Khan, A. (2016). The impact of examination system on teaching and learning at secondary level in Pakistan. *Journal of Education and Practice*, 7(14), 112–118.
- Lyons, L. (1998). Ethics in language testing. In K. Johnson & C. Johnson (Eds.), *Encyclopedic dictionary of applied linguistics* (pp. 126–127). Blackwell.
- Mahmood, K. (2012). *Examination systems in Pakistan: A study of prevailing practices*. National Book Foundation.

- Noble, J., & Smith, M. L. (1994). Old and new beliefs about measurement-driven instruction: The more things change, the more they stay the same. *Educational Measurement: Issues and Practice*, 13(4), 13–16. <https://doi.org/10.1111/j.1745-3992.1994.tb00451.x>
- Raza, W. (2009). Higher education testing of English language in Pakistan. *Pakistan Journal of Social Sciences*, 6(2), 87–94.
- Shamim, F. (2011). English as the language for development in Pakistan: Issues, challenges and possible solutions. In H. Coleman (Ed.), *Dreams and realities: Developing countries and the English language* (pp. 1–16). British Council.
- Xie, Q., & Jia, Q. (2025). A bibliometric review of language testing washback research (1993–2023). *Language Testing Research Review*, 19(1), 45–62.
- Yin, R. K. (2018). *Case study research and applications: Design and methods* (6th ed.). SAGE.