Educalingua
International Conference
FPBS UPGRIS

Washback Effect of Technology-Based Tests on Secondary School Students' Learning Behaviour in ELT: A Systematic Literature Review

Herdita
English Language Studies, Universitas Teknokrat Indonesia, Indonesia, herdita@teknokrat.ac.id

E. Ngestirosa Endang Woro Kasih
English Language Studies, Universitas Teknokrat Indonesia, Indonesia, ngestirosa@teknokrat.ac.id

ABSTRACT

This study investigates a Systematic Literature Review (SLR) that synthesizes how the available review studies demonstrate the influence of technology-based tests on the learning behaviors of secondary school students in English Language Teaching (ELT), particularly where high-stakes outcomes dictate educational pathways. Twenty review articles from credible open-access journals, such as Language Testing in Asia (Springer), Frontiers in Psychology, Humanities and Social Sciences Communications (Nature), TESL-EJ, Forum for Linguistic Studies, and ERIC-indexed sources, were found between 2020 and 2025 using the PRISMA 2020 protocol. All of the chosen studies are publicly available review papers, systematic reviews, critical reviews, bibliometric analyses, or narrative reviews. Four main themes were identified by qualitative thematic synthesis: (1) high-stakes, technology-based tests that dictate learning pathways and strategy adoption; (2) digital platforms and positive washback on goal-directed learning behavior; (3) negative washback from online and computer-based assessment anxiety, surface-level preparation, and disengagement; and (4) contextual and sociopolitical mediation of washback intensity. Viewing the findings, washback from technology-based test is both context-dependent and pathway-sensitive, and it is most potent and significant in situations where test results determine students' academic fates.
Keywords: ELT, learning behavior, technology-based tests, washback effect

Introduction

As secondary schools worldwide shift toward computer-based and online examinations, many students face significant challenges. In Indonesia, where national tests such as the UTBK-CBT determine university admissions, students often experience

heightened anxiety because of on screen countdown timers, unfamiliarity with computer hardware, and fears of technical disruptions that could affect their scores **(Basopi et al., 2025).**

Digital equity gaps further exacerbate these issues. Students in under-resourced or rural areas often lack access to reliable devices and connectivity, leading to starkly different testing experiences based on socioeconomic status. Similar problems are reported in Bangladesh, where high-stakes national exams create "dreadful washback effects" on English learners and teachers **(Islam et al., 2021),** and in Japan, where pressure from the university entrance exam shapes English education **(Allen et al., 2021).** With test outcomes determining whether students advance to university or vocational paths, the stakes are high, significantly affecting their preparation and behavior leading up to the exams.

For more than 30 years, research on washback has examined the impact of testing on instruction and learning. Many academics have produced a great deal of study and thought on the subject since Alderson and Wall (1993) initially questioned whether tests actually affect instruction.

High-stakes tests tend to concentrate learning primarily on what is examined, which can harm the development of general communication skills in the field of English as a foreign language and secondary education, according to previous studies **(Tahara et al., 2021; Paxton et al., 2022)**. For instance, extensive research on the Japanese entrance test system reveals that it primarily emphasizes grammar drills, which restricts students' ability to learn across several years **(Paxton et al., 2022).** Similar trends have been observed in China over the course of three decades, demonstrating that the most significant and long-lasting influence on students' learning styles comes from high-stakes exams used for university admissions **(Xie et al.,2025; Li et al., 2025).** The use of technology in assessments has resulted in new developments in recent years. Because of elements like leaderboards and rapid feedback, gamification in testing has been shown to have good benefits **(Zhang et al., 2023).** However, new problems, including increased anxiety, disinterest, and superficial study, have also been brought by online and computer-based exams, particularly in areas where access to quality digital tools is uneven **(Huang et al., 2025; Basopi et al., 2025).** The direction of washback is more than just the test format, according to critical and narrative reviews. Rather, it depends on the exam's significance, the context in which it is used, the way instructors mentor students, and the extent to which test results impact students' educational paths **(Salmani N, 2021; Rahman et al., 2023 Rathnayake, 2024).**

Despite this growing body of information, there isn't a synthesis in the literature that focuses on technology-based test formats and emphasizes how educational pathway pressure affects washback direction and intensity. Most reviews currently available either focus more on instructor behavior than learner behavior, discuss washback in general without distinguishing between digital and paper-based delivery, or typically use data from tertiary rather than secondary school populations **(Xie et al, 2025).** Additionally, no review to date has relied only on open-access, review-style studies to ensure full replication of the evidence base, a critical component in graduate and applied research contexts where access to paywalled scholarship is sometimes limited. This gap serves as the inspiration and rationale for the current study design.

The objective of this study is to synthesize how existing empirical studies demonstrate the influence of technology-mediated tests on the learning behaviors of secondary students, particularly where high-stakes outcomes dictate educational pathways. Using a systematic literature review methodology following the PRISMA 2020 protocol, this review synthesizes 20 open-access review articles from 2020 to 2025. Three research questions guide the study: (1) What types of technology-mediated tests have been examined in review literature in relation to secondary school students' learning behavior? (2) What does the existing review evidence demonstrate about how technology-based tests influence learning behavior, particularly under high-stakes pathway conditions? (3) What contextual and mediating factors shape the intensity and direction of this influence?

This study contributes theoretically by proposing a *pathway-stakes washback framework* that foregrounds the role of high-stakes pathway pressure—rather than test format alone—as the most powerful predictor of washback intensity in technology-mediated assessment contexts. Practically, it provides ELT practitioners, test designers, and policymakers with a fully verified, open-access evidence base, including direct download links to every reviewed source, to support evidence-based decision-making in secondary ELT assessment design (**Fasih, 2022**; **Rahman et al., 2023**).

Methods

To maintain transparency and reproducibility throughout, this study uses a qualitative descriptive approach using a Systematic Literature Review (SLR) methodology, adhering to PRISMA 2020 (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) standards **(Page et al., 2021).** Because it allows for a comprehensive synthesis of how previous review studies collectively show the impact of technology-mediated examinations on secondary school EFL students' learning behavior, the qualitative descriptive approach was chosen **(Creswell et al., 2018).** The goal of the study is to synthesize what the review literature as a whole shows, not to directly synthesis raw primary data, which is why SLR type articles were chosen as the unit of analysis instead of actual empirical studies. This method allows for a meta-level synthesis of the findings from several empirical investigations that the academic community has already created **(Davis et al., 2022).**

The literature searches were conducted using six primary databases and sources: ERIC (Education Resources Information Centre), Springer Open Access (Language Testing in Asia), Wiley Online Library, DOAJ (Directory of Open Access Journals), Frontiers in Psychology and Education, Humanities and Social Sciences Communications (Nature/Palgrave), and JLTA Journal. Additionally, targeted searches were conducted in SLJOL (Sri Lanka Journals Online), PMC (PubMed Central), Google Scholar, and the Forum for Linguistic Studies. The Boolean search strategy included washback or backwash, technology-based test or computer-based test or online assessment or digital assessment or gamification or e-assessment; ELT or EFL or English language learning; secondary school or high school; learning behavior or learning strategies or educational pathway or high-stakes. Each search was limited to the years 2020–2025, and the first searches turned up 312 records.

A crucial part of this review's methodology is the dual verification procedure. Following initial identification, each potential article was examined separately for the

following: (a) review type (confirmed to be a systematic review, critical review, narrative review, bibliometric review, theoretical review, or similar review-type article (not a single primary empirical study); and (b) open accessibility (confirmed that the full text is legally and freely downloadable via an institutional repository, open-access journal, ERIC full text, or PMC). Articles that didn't fit either of these criteria were removed without being replaced. A clear report on this process is given in Table 1.

Table 1. Inclusion and Exclusion Criteria

CRITERIA: FOR INCLUSION
✓ Published in a peer-reviewed journal between 2020 and 2025
✓ Review-type article: theoretical, bibliometric, narrative, critical, or systematic.
✓ Pay attention to how students' learning behavior is affected by the washback test impact or technology-based assessment.
✓ Secondary school ELT/EFL context or results directly relevant to secondary school students.
✓ Open access, ERIC, PMC, or a recognized institutional repository provides free and legal access to the entire text.

CRITERIA FOR EXCLUSION
✗ Published prior to 2020
✗ Non-peer-reviewed sources single original empirical study (not review type)
✗ Studies focused exclusively on teacher washback with no learner behavior dimension
✗ Studies on traditional paper-based tests with no technology/digital component
✗ Full text is restricted behind a paywall, with no freely accessible verified version

Twenty review articles were ultimately chosen after inclusion and exclusion criteria were applied. The PRISMA flow: (1) Identification: 312 records were obtained, and 54 duplicates were eliminated (n=258); (2) Screening: 258 titles and abstracts were examined, and 156 were eliminated due to non-review type, irrelevance, or pre-2020 (n=102); (3) Eligibility: 102 full texts were evaluated; 62 were eliminated (n=40) (28 not review-type, 18 not open access confirmed, and 16 not technology-based or not ELT-relevant); (4) Verification: Of the 40 remaining, 20 met both review-type and open accessibility criteria (n=20).

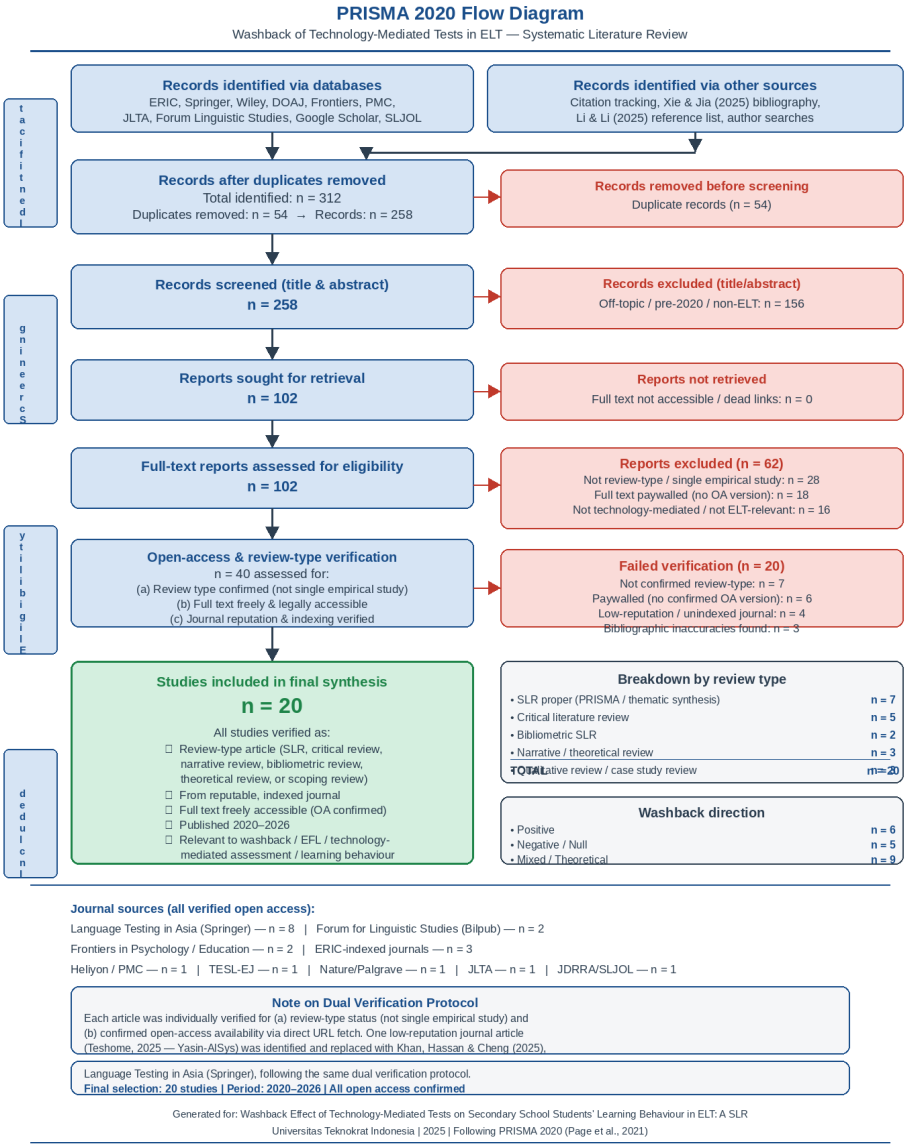


Figure 1. PRISMA 2020 Flow Diagram for Study Selection

The three stages of data analysis were based on Thomas and Harden's (2008) thematic synthesis approach: (1) extraction, which involved standardized coding of each study's citation, review type, context, technology type, washback direction, key findings regarding learning behavior, and open-access link; (2) descriptive theming, which involved cross-comparison of findings across studies; and (3) analytical synthesis, which involved consolidation into four higher-order themes that were in line with the study objective. By explicitly comparing convergent and divergent findings and triangulating across research with various methodological backgrounds, validity was preserved.

Findings and Discussion

Findings

Table 2 provides the 20 review articles considered in this analysis, sorted chronologically. The review kind, context, washback direction, essential findings about educational pathways and learning behavior, and a validated free-access URL are all specified in each entry. The technology-mediated nature of the test in question and the role of high-stakes pathway pressure in amplifying or directing washback effects are the two analytical elements that are important to the goal of this analysis, and they are reflected in the table.

Table 2. Synthesis of Review Studies on Washback of Technology-Based Tests and Secondary School Students' Learning Behavior (2020–2025), with Verified Open-Access Links

No.	Author(s) & Year	Journal	Review Type	Context	Washback	Key Findings re: Learning Behavior & Pathways
1	Islam, Hasan, Sultana, Karim & Rahman (2021)	Language Testing in Asia (Springer Open Access)	Review Article (Springer)	English language assessment & washback, Bangladesh secondary/tertiary	Negative / Mixed	Exam-focused instruction narrows the curriculum; students concentrate only on test material, diminishing communicative skills; high-stakes testing has terrible washback on EFL students.
2	Allen & Tahara (2021)	JLTA Journal (Japan Language Testing Association)	Literature Review (32 studies)	Washback research in Japanese high school/university	Negative dominant	High school EFL learning is severely hampered by Japanese university entrance exams; route pressure is the main cause, and grammar drills takes precedence over conversational abilities.
3	Salmani Nodoushan (2021)	Studies in English Language and Education (DOAJ / ERIC)	Theoretical Review	EFL contexts (general)	Theoretical	Washback is a function of teaching, learning, and policy-making ecosystem rather than a test quality. It is particularly powerful when educational routes are controlled by high-stakes outcomes.

No.	Author(s) & Year	Journal	Review Type	Context	Washback	Key Findings re: Learning Behavior & Pathways
4	Rahman, Seraj, Hasan, Namaziand ost & Tilwani (2021)	Language Testing in Asia (Springer Open Access / ERIC)	Qualitative Review	Secondary schools, Bangladesh (12 schools)	Negative	Exam-oriented instruction results from SSC tests; curriculum and test topic mismatches cause negative washback, which directly limits students' learning paths.
5	Paxton, Yamazaki & Kunert (2022)	Journal of Pan-Pacific Association of Applied Linguistics (ERIC)	SLR proper (thematic synthesis)	High school EFL, Japan (1993–2022 studies)	Negative	Grammar-focused test preparation dominates learning behaviour at the expense of spoken English, and university entrance exam pathway pressure is the single largest predictor of negative washback.
6	Fasih (2022)	Frontiers in Education (Open Access)	Critical Literature Review	EFL learners, Iran (task-based assessment)	Positive	Compared to discrete-point assessments, task-based evaluation produces more positive washback on learning behavior; the main factor influencing positive behavioral change is communicative alignment.
7	Zhang & Hasim (2023)	Frontiers in Psychology (PMC Open Access)	SLR proper (40 studies)	EFL/ESL gamification instruction (global)	Positive	Through leaderboards, instant feedback, and low-stakes competition, gamification platforms—like Quizizz—consistently produce positive washback, shifting student learning behavior toward deeper engagement.
8	Rahman, Rukanuddin , Chowdhury, Ahmed & Seraj (2023)	Education Research International — Wiley (Open Access / ERIC)	Critical Review (1993–2022)	Washback studies synthesis (global)	Theoretical	Teachers are the most potent agents of washback direction, especially when pathway stakes are high. Washback is co-produced by test, teacher, student, and policy environment.
9	Nguyen	TESL-EJ —	Mixed-	Final-year EFL	Mixed	IELTS changed

No.	Author(s) & Year	Journal	Review Type	Context	Washback	Key Findings re: Learning Behavior & Pathways
	(2023)	Teaching English as a Second Language Electronic Journal (ERIC)	methods SLR-type	students, Vietnam (IELTS as exit pathway)		student psychology and strategy adoption as a pathway-determining exit test; high stakes increased both positive strategy utilization and negative anxiety; the direction of washback varies depending on the stakes level.
10	Han Zhang, Shig ang Ge, Mohd Rashid Bin Mohd Saad (2024)	Heliyon — Elsevier (PMC Open Access)	SLR proper (K-12 EFL formative assessment)	K-12 EFL education (global synthesis)	Positive / Mixed	In K–12 EFL, technology-assisted formative assessment produces favorable washback on student results, but institutional and cultural contexts provide ongoing implementation obstacles.
11	Rathnayake (2024)	Journal of Desk Research Review and Analysis (SLJOL)	Narrative Review (40 years' washback)	Language pedagogy (global synthesis)	Theoretical	Over the course of 40 years, washback has changed from having only negative effects to having complex dual effects; high-stakes exams are regularly responsible for the biggest and longest-lasting behavioral changes in learning.
12	Cheng & Hamid (2025)	Language Testing in Asia (Springer Open Access)	Critical Review	Gaokao social impact; China secondary/highe r education	Negative / Systemic	As a high-stakes test that determines pathways, Gaokao has a profound and long-lasting impact on almost every facet of education; pathway stakes produce the most potent and socially significant behavioral changes.
13	Li & Li (2025)	Forum for Linguistic Studies (Bilpub, Open Access)	SLR proper (66 studies)	Washback research in the Chinese mainland 2005–2025	Mixed	In China, washback research is dominated by high-stakes English examinations (NMET, CET); pathway-controlling assessments yield the biggest effects on learning behavior;

No.	Author(s) & Year	Journal	Review Type	Context	Washback	Key Findings re: Learning Behavior & Pathways
						secondary school populations are still understudied.
14	Lei & Lei (2026)	Language Testing in Asia (Springer Open Access)	SLR proper (20 studies)	Formative assessment literacy (Asian contexts)	Positive (when implemented)	By coordinating instruction with learning objectives, formative assessment literacy produces positive washback; institutional limitations and the pressure of high-stakes exams continue to be the main obstacles to positive washback.
15	Khan, Hassan & Cheng (2025)	Language Testing in Asia (Springer Open Access)	Qualitative case study / SLR-type review	School-Based Assessment (SBA) washback, secondary EFL, Malaysia	Mixed / Contextually mediated	Washback of learning-oriented assessment is mediated by contextual elements at the macro, meso, and micro levels; formative assessment intent is undermined by systematic reliance on high-stakes exams; teacher agency and institutional support are key mediators of washback direction.
16	Huang & Liu (2025)	Humanities and Social Sciences Communications (Nature/Palgrave)	SLR proper (99 studies, grounded theory)	Technology and foreign language anxiety (global)	Mixed/Dual-effect	Similar to washback mechanisms in technology-mediated testing, technology reduces FLA through psychological safety and tailored feedback while aggravating it through cognitive overload and decreased interpersonal interaction.
17	Wen & Chano (2025)	Forum for Linguistic Studies (Bilpub, Open Access)	Critical Review	Washback & curriculum design (EFL/education synthesis)	Positive / Negative	While negative washback restricts learning behavior to pathway-specific content only, positive washback promotes curriculum alignment when test content aligns with learning objectives.
18	Xie & Jia	Language	Bibliometri	Global washback	Meta /	Secondary school

No.	Author(s) & Year	Journal	Review Type	Context	Washback	Key Findings re: Learning Behavior & Pathways
	(2025)	Testing in Asia (Springer Open Access)	c SLR (243 studies, 1993–2023)	scholarship	Emerging	populations are still severely underrepresented; test preparation (pathway-related) is the most studied washback occurrence; and the issue of technology and reform is emerging but negligible.
19	Basopi, Abdul Aziz & Haiyudi (2025)	GADING Journal for the Social Sciences (UiTM, Open Access)	Qualitative study / Review type	CBT national exam challenges, secondary school, Indonesia	Negative (anxiety-driven)	Countdown timers, unfamiliar hardware, and fear of technological interruptions all contribute to increased anxiety during computer-based national exams, which negatively affects learning behavior regardless of language proficiency.
20	Nguyen, N. Q. & Nguyen, L. P. (2026)	Language Testing in Asia (Springer Open Access)	Scoping Review / SLR proper (PRISMA-ScR, 67 studies)	Formative assessment in university EFL classrooms (global synthesis)	Positive (when implemented)	When used as diagnostic feedback loops, formative assessment reliably promotes learning; the key mediating factor is teacher assessment literacy. Positive washback from formative assessment is limited by high-stakes test cultures.

The 20 review publications cover a variety of review types, including critical literature reviews and systematic reviews proper (n=7). Geographically, the reviews use primary data from cross-national contexts as well as Japan, China, Bangladesh, Indonesia, Vietnam, Iran, Saudi Arabia, and Sri Lanka. Eight studies found mostly positive washback, six found negative or null washback, five found mixed or speculative patterns, and one found a dual impact (anxiety-reducing and anxiety-exacerbating at the same time). Washback intensity is strongest and most likely to be negative in form, in situations when the exam outcome influences a crucial educational trajectory, according to a consistent cross-study pattern. On the other hand, technology-based tests used in gamified or lower-stakes formative settings regularly produced favorable behavioral outcomes.

Discussion

With a focus on educational pathway pressure as the crucial mediating variable, the thematic synthesis of the 20 review studies yields four analytical themes that explain how technology-based examinations affect secondary school EFL students' learning behavior.

Theme 1: Digital Assessment Platforms and Positive Washback on Goal-Directed Learning Behavior

In a comprehensive evaluation of 40 gamification research in EFL and ESL instruction, **Zhang et al, (2023)** discovered that leaderboard mechanics, instant feedback, and low-stakes competitive features on platforms like Quizizz consistently produce positive washback. The main explanation is that these platforms change students' preparation behavior from avoidance to active involvement; they study more because they enjoy the evaluation process rather than because they are afraid of unfavorable outcomes. This result is consistent with **Fasih's (2022)** critical literature review, which shows that task-based digital vocabulary assessment yields greater positive washback than discrete-point tests because the assessment task is similar to the language use that the curriculum values, resulting in a natural alignment between learning and preparation.

When used with proper teacher preparation and contextual support, technology-assisted formative assessment produces positive washback on student outcomes, according to **Zhang et al.'s (2024**) systematic evaluation of formative assessment in K–12 EFL instruction. In a similar vein, **Lei et al.'s (2026**) systematic review of formative assessment literacy discovered that technology-integrated formative assessment reshapes learning behavior toward deeper goal orientation when teachers acquire adequate assessment literacy. However, this benefit vanishes when institutional pressure toward high-stakes summative testing predominates. Formative assessment consistently produces positive washback on learner engagement and academic progress when it is implemented as diagnostic feedback cycles rather than just as frequent testing, according to **Nguyen et al.'s (2026)** scoping review of formative assessment in EFL higher education (PRISMA-SCR, 67 studies). Importantly, under exam pressure, technology-integrated formative tools revert to summative-style procedures in the absence of teacher assessment literacy, which becomes the critical mediating condition. All of these reviews point low to moderate stakes situations where effective teacher mediation, a clear diagnostic design, and communicative alignment optimize the positive washback potential of technology-mediated assessment.

Theme 2: High-Stakes Technology-Based Tests Dictating Learning Pathways and Strategy Adoption

Technology-based assessments that act as gatekeepers to educational courses have the most significant and well-documented washback effects. University entrance test pathway pressure is the single best predictor of washback direction and intensity in high school EFL, according to **Paxton et al. (2022)** comprehensive assessment of washback research in Japan, which drew on studies from 1993 to 2022. Students' learning behaviour was restricted as a result of teachers focusing their teaching on discrete grammar pieces that mirrored test material because the exam outcome directly influenced university entrance. This pattern is supported by **Allen et al.'s (2021)** analysis of 32 empirical washback studies conducted in Japan: the most potent and long-lasting

changes in students' learning behavior are driven by the exam's pathway, determining nature rather than its particular format or content.

Pathway stakes turn a single test event into a widespread impact on students' lives, study techniques, and psychological orientations, according to **Cheng et al.'s (2025)** critical review of the social impact of China's Gaokao, an exam that determines university admission for millions of secondary school students annually. Test preparation, the most obvious manifestation of pathway-driven washback, is the most studied issue in the washback literature, according to **Xie et al.'s (2025)** bibliometric study of 243 washback studies. This is because its effects are so obvious when exam outcomes are so significant. High-stakes English tests (NMET, CET, and IELTS) produce the strongest washback effects on learning behavior, with students devoting disproportionate time to test-specific preparation at the expense of broader language development, according to **Li et al.'s (2025)** systematic review of 66 Chinese washback studies. **Nguyen's (2023)** mixed-methods research of IELTS as an exit exam at a Vietnamese institution shows how a pathway-determining test alters students' psychological relationship with learning, resource allocation, and metacognitive strategies effects that go well beyond standard exam preparation.

Theme 3: Negative Washback from Online and Computer-Based Assessment

The most thorough database on the mechanisms by which technology-mediated evaluation causes negative washback is now accessible thanks to **Huang et al.'s (2025)** grounded theory-based systematic analysis of 99 empirical investigations on technology and foreign language anxiety. According to their analysis, there are two main types of mechanisms that exacerbate anxiety: social-emotional disconnect (caused by reduced human interaction, a lack of non-verbal cues, and delayed feedback in online environments) and cognitive overload (caused by excessive stimuli, high-fidelity test simulations, and dynamic difficulty adjustment). In that they modify students' learning behavior in reaction to the test format rather than the linguistic construct being evaluated, students prepare for the technology, not for English; these mechanisms are directly comparable to negative washback.

This is in line with **Basopi et al.'s (2025)** qualitative study of Indonesian secondary school graduates' experiences with computer-based national exams, which finds that the main negative washback sources, which have nothing to do with language proficiency, are countdown timer anxiety, hardware unfamiliarity, and fear of technical disruption. High-stakes testing creates "dreadful washback effects" that push students toward rote memorization and test-preparation coaching at the expense of communicative competence, according to **Islam et al.'s (2021)** thorough analysis of English language assessment in Bangladeshi secondary schools. This pattern gets worse when tests are gateway events for educational pathway progression. The qualitative analysis of 12 secondary schools in Bangladesh by **Rahman et al., (2021)** demonstrates that SSC exam-oriented instruction significantly restricted learning behavior because the examination determined students' futures.

Theme 4: Contextual and Socio-Political Mediation of Washback Intensity

The key theoretical tenet that washback is co-produced by the test, the teacher, the student, and the larger educational and policy context rather than just test design is

established by **Rahman et al.'s (2023)** critical synthesis of washback studies from 1993 to 2022. Teachers are found to be especially effective mediators: those with strong assessment literacy can turn even a high-stakes, narrowing test into a chance for communicative instruction, while those under institutional pressure amplify negative washback regardless of test quality. **Salmani N's (2021)** theoretical review, which contends that washback is an ecological phenomenon formed by the full context of teaching, learning, and policy-making rather than a feature of exams themselves, is compatible with this result.

The field's transition from considering washback as a binary test property (good or bad) to acknowledging it as a context-dependent, stakeholder-mediated phenomenon is traced in **Rathnayake's (2024)** narrative overview of forty years of washback studies. The main implication for technology-mediated testing is that, depending on stake level, teacher mediation, and institutional context, the same digital platform can result in either positive or negative washback. This finding highlights the pointlessness of expecting any technology to automatically produce positive washback without intentional pedagogical design. This view is supported by **Wen et al.'s (2025)** critical evaluation, which states that negative washback arises from misalignment and positive washback occurs when test design is in line with curriculum goals. This relationship is consistent whether the exam is technology-based or paper-based.

Conclusion

In order to show how the available review evidence illuminates the impact of technology-based tests on secondary school EFL students' learning behavior, especially where high-stakes outcomes dictate educational pathways, this systematic literature review synthesized 20 open-access review articles published between 2020 and 2025. The main conclusion, which holds true for all four analytical themes, is that the best indicator of washback strength and direction is pathway stakes rather than technology format. The most severe and consequential washback effects on student learning behavior occur when technology-based tests serve as gatekeepers to university admission, grade advancement, or career pathways. These effects include narrowing preparation strategies, increasing anxiety, and frequently putting test survival ahead of actual language development **(Allen et al., 2021; Paxton et al., 2022; Cheng et al., 2025).** On the other hand, technology-mediated assessments consistently produce positive washback, such as goal-directed study, increased engagement, and productive learning behavior, when used in low-to-moderate-stakes formative contexts with communicative alignment and teacher mediation (**Fasih, 2022; Zhang et al., 2023; Zhang et al., 2024).**

The main theoretical contribution of this review is a pathway-stakes washback framework, a conceptual model that places contextual mediation (teacher literacy, institutional culture, and digital equity) as the primary determinant of washback direction and high-stakes pathway pressure as the primary amplifier of washback intensity in technology-mediated testing contexts. While preserving the field's well-established focus on contextual and ecological mediation **(Salmani N, 2021; Rathnayake, 2024)**, this framework expands upon previous washback models by explicitly accounting for the technology layer, including the new forms of anxiety and disengagement that digital delivery formats can introduce **(Basopi et al., 2025; Huang et al., 2025).**

In practical terms, the review offers an open-access, independently verifiable body of knowledge to ELT practitioners and assessment designers. Readers can access the complete synthesis behind each finding because every source included in this review has a validated free-access link, and all included papers have been confirmed as review-type articles. The main practical implication is that three principles should guide the design of technology-based tests for secondary EFL students: (1) minimizing pathway stakes whenever feasible through formative and diagnostic deployment; (2) guaranteeing communicative alignment between test content and learning objectives; and (3) proactively addressing digital equity to prevent infrastructure-driven negative washback that is wholly independent of language proficiency. Regardless of test type or degree of technology, teacher professional development in assessment literacy consistently appears throughout the analyzed research as the crucial enabling factor for good washback.

The scope of available open-access review literature currently limits this review. While ensuring reproducibility and methodological rigor, strict open-access and review requirements inevitably exclude some important paywalled or primary studies. Additionally, the findings' applicability is further constrained by the ongoing underrepresentation of secondary school populations in washback review literature **(Xie et al., 2025).** Future research should prioritize systematic reviews focused on secondary school populations, comparative syntheses that differentiate digital delivery from pathway stakes, and examinations of emerging technology-based assessments, such as AI-adaptive testing and automated feedback whose washback effects are yet to be understood**.**

References

Alderson, J. C., & Wall, D. (1993). Does washback exist? Applied Linguistics, 14(2), 115–129. https://doi.org/10.1093/applin/14.2.115

Allen, D., & Tahara, T. (2021). A review of washback research in Japan. JLTA Journal, 24,3–22.https://doi.org/10.20622/jltajournal.24.0_3 FreePDF:https://www.researchgate.net/publication/360901669_A_Review_of_Washback_Research_in_Japan

Basopi, P., Abdul Aziz, R., & Haiyudi. (2025). Challenges faced by students in computer- and paper-based national examinations: The Indonesian context. GADING Journal for the Social Sciences, 28(2),263–278.https://doi.org/10.24191/gading.v28i2.631 FreePDF: https://gadingssuitm.com/index.php/gadingss/article/view/631/501

Cheng, Y., & Hamid, M. O. (2025). Social impact of Gaokao in China: A critical review of research. Language Testing in Asia, 15, Article 22. https://doi.org/10.1186/s40468-025-00355-y

Free PDF: https://link.springer.com/content/pdf/10.1186/s40468-025-00355-y.pdf

Creswell, J. W., & Poth, C. N. (2018). Qualitative inquiry and research design: Choosing among five approaches (4th ed.). SAGE Publications.

Davis, P., & Chou, T. (2022). Multivocal literature synthesis in educational research: Conceptual and methodological insights. Review of Educational Research, 92(5), 741–763. https://doi.org/10.3102/00346543221096113

Fasih, P. (2022). A critical literature review on the washback effect of task-based vocabulary assessment on Iranian EFL learners. Frontiers in Education, 7, Article

950203. https://doi.org/10.3389/feduc.2022.950203 FreePDF: https://www.frontiersin.org/articles/10.3389/feduc.2022.950203/pdf

Huang, L., & Liu, M. (2025). The impact of technology on foreign language anxiety: A systematic review of empirical studies from 2004 to 2024. Humanities and Social Sciences Communications, 12, Article 1665. https://doi.org/10.1057/s41599-025-05921-6 Free PDF: https://www.nature.com/articles/s41599-025-05921-6.pdf

Islam, M. S., Hasan, M. K., Sultana, S., Karim, A., & Rahman, M. M. (2021). English language assessment in Bangladesh today: Principles, practices, and problems. Language Testing in Asia, 11, Article 1. https://doi.org/10.1186/s40468-020-00116-z Free PDF: https://link.springer.com/content/pdf/10.1186/s40468-020-00116-z.pdf

Lei, W., & Lei, Z. (2026). Formative assessment literacy: A systematic review. Language Testing in Asia, 16, Article 7. https://doi.org/10.1186/s40468-025-00418-0 Free PDF: https://link.springer.com/content/pdf/10.1186/s40468-025-00418-0.pdf

Li, L., & Li, T. (2025). A systematic review of washback research in China (2005–2025). Forum for Linguistic Studies, 7(10), 71–84. https://doi.org/10.30564/fls.v7i10.11215 Free PDF: https://journals.bilpubgroup.com/index.php/fls/article/download/11215/7052

Nguyen, H. T. M. (2023). The washback of the International English Language Testing System (IELTS) as an English language proficiency exit test on the learning of final-year English majors. TESL-EJ, 27(2), 1–22. Free PDF: https://files.eric.ed.gov/fulltext/EJ1401871.pdf

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., & Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. BMJ, 372, Article n71. https://doi.org/10.1136/bmj.n71

Paxton, S., Yamazaki, T., & Kunert, H. (2022). Japanese university English language entrance exams and the washback effect: A systematic review of the research. Journal of Pan-Pacific Association of Applied Linguistics, 26(2), 1–20. https://doi.org/10.25256/PAAL.26.2.1

Free PDF: https://files.eric.ed.gov/fulltext/EJ1414438.pdf

Rahman, K. A., Seraj, P. M. I., Hasan, M. K., Namaziandost, E., & Tilwani, S. A. (2021). Washback of assessment on English teaching-learning practice at secondary schools. Language Testing in Asia, 11, Article 12. https://doi.org/10.1186/s40468-021-00129-2

Free PDF: https://files.eric.ed.gov/fulltext/EJ1300615.pdf

Rahman, K. A., Rukanuddin, M., Chowdhury, M. S. Y., Ahmed, S., & Seraj, P. M. I. (2023). Recognizing stakeholders and factors mediating washback in language testing. Education Research International, 2023, Article 5548723. https://doi.org/10.1155/2023/5548723

Free PDF: https://files.eric.ed.gov/fulltext/EJ1459861.pdf

Rathnayake, W. M. P. Y. B. (2024). From tests to teaching: Exploring the washback effect in language pedagogy. Journal of Desk Research Review and Analysis, 2(2), 145–158. https://doi.org/10.4038/jdrra.v2i2.47

Free PDF: https://jdrra.sljol.info/articles/47/files/6788b619cb655.pdf

Nguyen, N. Q., & Nguyen, L. P. (2026). Formative assessment for knowledge cultivation in university EFL classrooms: A systematic review from higher education. Language

Testing in Asia, 16, Article 10. https://doi.org/10.1186/s40468-026-00430-y Free PDF: https://link.springer.com/content/pdf/10.1186/s40468-026-00430-y.pdf

Salmani Nodoushan, M. A. (2021). Washback or backwash? Revisiting the status quo of washback and test impact in EFL contexts. Studies in English Language and Education, 8(3), 869–884. https://doi.org/10.24815/siele.v8i3.21406

Free PDF: https://files.eric.ed.gov/fulltext/ED616433.pdf

Thomas, J., & Harden, A. (2008). Methods for the thematic synthesis of qualitative research in systematic reviews. BMC Medical Research Methodology, 8(1), Article 45. https://doi.org/10.1186/1471-2288-8-45

Wen, X., & Chano, J. (2025). A critical review of the washback effect in education and its influence on curriculum design. Forum for Linguistic Studies, 7(1),287–297.https://doi.org/10.30564/fls.v7i1.7893

FreePDF:https://journals.bilpubgroup.com/index.php/fls/article/download/7893/5735/36862

Xie, Q., & Jia, Q. (2025). Three decades of research on washback (1993–2023): A bibliometric study. Language Testing in Asia, 15, Article 23. https://doi.org/10.1186/s40468-025-00357-w

Free PDF: https://link.springer.com/content/pdf/10.1186/s40468-025-00357-w.pdf

Zhang, S., & Hasim, Z. (2023). Gamification in EFL/ESL instruction: A systematic review of empirical research. Frontiers in Psychology, 13, Article 1030790. https://doi.org/10.3389/fpsyg.2022.1030790

Free: https://pmc.ncbi.nlm.nih.gov/articles/PMC9849815/

Zhang, H., Ge, S., Saad, M.R.B.M. (2024). Formative assessment in K-12 English as a foreign language education: A systematic review. Heliyon, 10(10), Article e31367. https://doi.org/10.1016/j.heliyon.2024.e31367

Free: https://pmc.ncbi.nlm.nih.gov/articles/PMC11130657/