



Language Assessment Knowledge and Practice among Libyan School English Teachers: A Critical Review of the Literature

Author name: Dr. Ahmed A. M. Ahmed

Affiliation: Higher Institute of Medical Sciences and Technologies - Alshatti

correspondence details: a.ahmed_8283@yahoo.com

Received: 29/6/2026 Revised: 20/7/2026 Accepted: 28/7/2026 Publication: 10/8/2026

Abstract

Teachers' knowledge of language assessment is widely regarded as a determinant of the quality of classroom testing, feedback, and, ultimately, learning outcomes. Yet the great majority of research on language assessment literacy (LAL) has been conducted in higher-education or Western contexts, leaving the assessment knowledge and practice of school-level English as a Foreign Language (EFL) teachers in fragile and transitional education systems comparatively under-examined. This paper critically reviews the body of research, published between 2010 and 2026, that addresses the language assessment knowledge, beliefs, and practices of Libyan school English teachers. Seventeen primary are synthesised thematically around five recurring concerns: teachers' knowledge and practice of oral/speaking assessment; the relationship between assessment beliefs and classroom practice; the washback of high-stakes public examinations on teaching; the quality and curriculum-alignment of centrally set tests; and the uptake of alternative and dynamic assessment. The review argues that the Libyan literature converges on a picture of persistent negative washback and a curriculum–examination misalignment inherited from the 2000 curriculum reform, but diverges sharply on whether the underlying cause is best explained by individual teachers' limited assessment literacy or by structural features of what one commentator has termed Libya's "testing-laden culture." The review closes by identifying methodological weaknesses in the existing evidence base — heavy reliance on small, single-region qualitative samples and self-report instruments, and proposes a research agenda centred on validated LAL measurement, classroom-observation triangulation, and intervention studies at the national school level.

Keywords: language assessment literacy; teacher knowledge; washback; Libya; EFL; secondary school assessment

1. Introduction

Assessment is not incidental to language teaching; it is one of the principal mechanisms through which curricular intentions are translated — or distorted — into classroom practice (Alderson & Wall, 1993; Bachman & Palmer, 2010). Where teachers possess comprehensive knowledge of assessment purposes, principles, and techniques, classroom testing is more likely to support learning; where that knowledge is insufficient, or where systemic incentives override it, assessment tends to narrow instruction toward test preparation regardless of curricular aims (Cheng, 2004). This tension is especially visible in contexts where a single high-stakes public examination governs progression through, or exit from, school education.

Libya presents an unusually instructive, and unusually under-studied, case. The 2000 English-language curriculum reform introduced a learner-centred, communicative syllabus, yet the accompanying public examination system was not simultaneously redesigned, producing a durable mismatch between what the curriculum asks teachers to do and what the national high-stakes examination as the Secondary Education Certificate Examination (SECE), rewards (Orafi & Borg, 2009). This mismatch has been compounded, by prolonged political instability that has disrupted teacher training infrastructure, supervision, and continuing professional development across the country. Within this setting, a body of doctoral and journal-level research — produced largely by Libyan scholars completing postgraduate study abroad — has accumulated around questions of what Libyan school English teachers know about assessment, what they believe about it, and how they actually assess their students.

This literature has not, to date, been synthesised. Individual studies address speaking assessment, general assessment beliefs, examination washback, test design quality, and alternative assessment as largely separate lines of inquiry. The present paper addresses this gap by critically reviewing seventeen primary sources on Libyan school-level English language assessment, asking three questions: (1) What does the existing literature suggest about Libyan school EFL teachers' knowledge and practice of language assessment? (2) To what extent, and how, does the literature explain observed patterns of practice — as a matter of individual teacher knowledge, or as a matter of systemic and policy conditions? (3) What methodological and substantive gaps remain? The

review proceeds by first situating the discussion within the wider language assessment literacy (LAL) framework, then describing the review method, before presenting a thematic critical synthesis and a methodological critique that together motivate a future research agenda.

2. Conceptual Framework: Language Assessment Knowledge and Literacy

The construct most often invoked to describe what teachers need to know about assessment is language assessment literacy (LAL): teachers' capacity to design, select, administer, interpret, and act upon the results of language assessments (Davies, 2008; Fulcher, 2012; Inbar-Lourie, 2008). Taylor's (2013) framework, probably the most widely applied in the EFL literature, distinguishes knowledge of theory (validity, reliability, practicality), technical skills (item and task design, scoring, statistics), and principles governing the appropriate use of assessment information; Giraldo (2020) later argued for supplementing this with context-specific and sociocultural dimensions, a move with particular relevance for a setting such as Libya, where assessment cannot be understood apart from the political economy of high-stakes national examinations.

The Libyan literature under review sits somewhat inconsistent within this mainstream LAL framework. Relatively few of the studies reviewed here set out explicitly to measure LAL as a defined construct (in the manner of, for example, the Language Assessment Literacy Questionnaire tradition built on Taylor, Fulcher, Hasselgreen et al., and Kremmel and Harding's instruments); most instead approach teacher assessment knowledge indirectly, through studies of beliefs, practices, or the washback of a specific examination. This is a substantive point for a critical review to register at the outset: "knowledge," "belief," "practice," and "washback" are treated in this literature as closely related but analytically distinct constructs, and conflating them — as some individual studies risk doing — obscures rather than clarifies what is actually being claimed. One contribution in the set reviewed here, Elmahjoub's critical review of the "testing-laden culture" of English language testing, is itself a conceptual rather than empirical piece, and its framing — asking whether Libyan teachers' assessment conduct reflects a lack of assessment savviness or an inherited policy tradition — supplies the organising question for the synthesis in Section 5.

3. Review Method

This is a critical, rather than systematic, literature review: its purpose is to build an argument about the state and coherence of a body of research, not to exhaustively map a construct through a fully reproducible search protocol. That said, transparency about scope strengthens rather than weakens a critical review, and the following parameters were applied. Sources were identified through targeted searching of institutional repositories mostly from the UK university theses), Google Scholar, and Libyan and regional journal platforms, using combinations of the terms "Libya," "EFL," "assessment," "testing," "washback," and "school teachers." Seventeen primary sources published or completed between 2014 and 2026 were selected for inclusion: fourteen address the Libyan school EFL context directly; a further two (Rahman et al., 2021, on Bangladesh; Al-Jaro & Tayeb, 2026) were retained as international comparators because they examine washback constructs directly relevant to the Libyan findings and allow the review to assess whether the Libyan pattern is context-specific or part of a wider pattern in high-stakes EFL testing contexts.

Sources were excluded if they addressed Libyan EFL teaching, curriculum, or teacher belief more generally without a substantive assessment or testing component (a large adjacent literature exists on Libyan EFL teacher beliefs, communicative language teaching implementation, and textbook use, which is referenced only where it directly informs the assessment discussion). Two limitations of this method should be stated plainly. First, several included sources are published doctoral theses or appear in Libyan and regional journals not indexed in Scopus or Web of Science; access to full text was limited, a constraint common to reviews of this literature given how much Libyan assessment research circulates outside mainstream indexing, Second, no Arabic-medium sources were reviewed in given that at least three of the seventeen sources are published in Arabic-language Libyan journals, this is a genuine limitation of the present review's linguistic reach that a fully systematic follow-up should address.

Language Assessment Knowledge and Practice among Libyan School English Teachers: A Critical Review of the Literature

Table 1. Summary of Studies Included in the Review (n = 17)

Study (Author, Year)	Context / Setting	Method	Key Focus / Finding
Grada (2014)	Libyan secondary schools; novice EFL teachers (Exeter thesis)	Qualitative; socio-cultural framework (interviews / classroom-based inquiry)	Novice teachers' speaking-assessment knowledge and practice are shaped by the school's wider activity system (colleagues, supervisors, exam pressure), not by individual knowledge alone.
El-Taher Al-Fourganee (2018)	Libyan secondary schools (Sunderland thesis)	Qualitative doctoral study	Examines how secondary teachers assess students' oral communication; documents inconsistent, underdeveloped speaking-assessment practice.
Nimehchisalem & Abdalla (2020)	Libyan secondary EFL teachers	Survey / mixed-methods (beliefs and practice measures)	Identifies a mismatch between teachers' stated assessment beliefs and their actual classroom assessment practice.
Elmahjoub (2023)	Libyan EFL teachers; dynamic assessment	Qualitative conceptions study	Teachers' limited uptake of dynamic assessment reflects a tension between conceptual understanding ("savviness") and an entrenched summative testing tradition.
Elmahjoub (n.d.)	Libya; English language testing culture	Conceptual / critical review (theoretical)	Frames Libya's "testing-laden culture" as a structural driver of assessment conduct; supplies the field's central savviness-vs-tradition question.
Ahmed (2019)	Libyan secondary school education (Huddersfield thesis)	Mixed-methods; Weir's socio-cognitive validity framework	SECE exam-preparation strategies raised exam scores without improving broader English learning outcomes; test format and exam-related cheating threaten score validity.
Onaiba (2014)	Libyan secondary EFL (Leicester thesis)	Doctoral study; perception/practice data on a revised public exam	Investigates washback of a revised EFL public examination on teachers' instructional practices, materials, and curriculum use.
Abdulhamid (2019)	Libyan EFL context (Carleton thesis)	Mixed-methods	Reframes washback as a function of curriculum-examination alignment rather than a fixed property of the exam itself.
Ali (2024)	Libyan ninth-grade EFL learners	Empirical washback study	Extends washback evidence to a lower-secondary population, reinforcing the exam-driven teaching pattern found at SECE level.
Elbakay, Zaglom, Elkut & Farid (2024)	Libyan secondary school teachers	Case study	Documents how high-stakes examinations directly shape and constrain teachers' instructional practices.
Masoud (2017)	Libyan General Certificate of	Test evaluation / document analysis	Evaluates the technical quality of GCSE English-language tests, relevant to their validity and fairness.

Language Assessment Knowledge and Practice among Libyan School English Teachers: A Critical Review of the Literature

Study (Author, Year)	Context / Setting	Method	Key Focus / Finding
	Secondary Education English tests		
Onaiba (2024)	Libyan secondary exam and curriculum documents	Evaluative document-analysis study	Assesses whether examination aims and content actually reflect stated curricular aims; evidence of misalignment from the instrument-design side.
Saleh & Megaiah (2025)	Two secondary schools, Al-Bayda City, Libya	Test-development case study with teachers and students	Designs and trials a new English-language secondary-certificate test, moving from diagnosis toward a concrete testing intervention.
Onaiba & Abudagga (2024)	Libyan EFL students	Quantitative outcome study (performance and attitude measures)	Alternative assessment methods affect both students' test performance and their attitudes toward assessment, offering rare outcome-level evidence.
Omeman (2019)	Libyan higher education sector	Doctoral qualitative study	Teachers' and students' views on assessment's impact on teaching/learning at university level; comparator to school-level findings.
Al-Jaro & Tayeb (2026)*	Non-Libyan EFL context (international comparator)	Classroom observation study	Directly observes how high-stakes tests cause teachers to substitute curriculum content with test-driven content — a design largely absent from the Libyan literature.
Rahman, Seraj, Hasan, Namaziandost & Tilwani (2021)*	Bangladeshi secondary schools (international comparator)	Mixed-methods washback study	Finds negative washback from curriculum–examination misalignment and explicitly cites Orafi and Borg's Libyan findings as a parallel case.

Note. Studies marked with an asterisk () are international comparator studies (Bangladesh; a non-Libyan EFL setting) included for benchmarking rather than direct evidence about the Libyan context. "Context/Setting" reflects the population and country studied; "Method" and "Key Focus/Finding" are drawn from available abstracts, theses records, and citing literature; primary full texts should be consulted to verify specific results before this table is used in a submitted manuscript, particularly for sources accessed only via title/metadata (see Section 3).*

4. Thematic Critical Review

4.1 Teachers' Knowledge and Practice of Speaking Assessment

Two studies focus specifically on the assessment of oral communication, arguably the skill most vulnerable to being under-assessed in examination-dominated systems because speaking is difficult to test reliably at scale. Grada (2014), working from a socio-cultural perspective at the University of Exeter, investigated Libyan EFL novice secondary school teachers' current knowledge and practice of speaking assessment, framing novice teachers' assessment conduct as

shaped by the wider activity system of the school — colleagues, supervisors, and the examination system — rather than as a purely individual cognitive matter. This socio-cultural framing is analytically important: it locates the explanation for gaps in speaking-assessment knowledge in the conditions novice teachers are inducted into, rather than treating limited knowledge as a deficit to be remedied by training alone. El-Taher Al-Fourganee (2018), in a doctoral study at the University of Sunderland, examined how secondary school Libyan teachers assess students' oral communication, complementing Grada's novice-teacher focus with a broader look at assessment practice.

Read together, these two studies suggest that speaking assessment is a persistently weak point in Libyan school EFL assessment — consistent with a wider regional pattern in which oral skills are marginalised in test-driven systems (as documented, for instance, in the broader Libyan EFL speaking-instruction literature). What neither study appears to resolve, however, is the relative contribution of teachers' own assessment knowledge versus external constraints (large classes, absence of speaking components in the SECE, limited teacher training in oral rubric design) to this pattern — precisely the knowledge-versus-context tension that recurs throughout the wider literature reviewed here.

4.2 Assessment Beliefs and the Knowledge–Practice Gap

Nimehchisalem and Abdalla (2020) offer the most direct empirical treatment of Libyan secondary school English teachers' assessment beliefs and practices, and their study has since been cited by researchers elsewhere (for example, in university-level LAL research in other contexts) as evidence that teachers' assessment beliefs and their actual classroom practices frequently diverge — a mismatch pattern echoed in formative-assessment research beyond Libya. This finding is consequential for how the whole literature set should be read: it cautions against inferring teachers' actual classroom assessment conduct directly from what they report believing or knowing, a methodological point that bears on several of the interview- and questionnaire-based studies discussed elsewhere in this review.

Elmahjoub's (2023) study of Libyan EFL teachers' conceptions of dynamic assessment extends this beliefs-oriented strand to a specific alternative-assessment construct, asking whether teachers' limited engagement with dynamic, mediation-based assessment approaches reflects a genuine lack

of "assessment savviness" or, instead, the pull of an entrenched policy tradition that continues to reward summative, examination-style testing regardless of individual teachers' conceptual grasp of formative alternatives. This question — savviness versus tradition — is arguably the single most important interpretive division in the literature reviewed here, and it recurs, in different guises, in the washback studies reviewed next.

4.3 High-Stakes Examinations and Washback on Teaching

The largest single cluster of studies in this literature concerns the washback of Libya's public examination system on classroom teaching, a pattern consistent with the wider washback literature's long-standing interest in how high-stakes tests shape instruction (Alderson & Wall, 1993; Cheng, 2004). Ahmed's (2019) doctoral thesis, examining English language teaching and learning in Libyan secondary school education through Weir's socio-cognitive validity framework, found that the teachers adopted strategies that improved students' examination performance on the public examination of SECE without a corresponding improvement in their broader English learning outcomes, and identified test content, format, and a social tolerance of cheating as threats to score validity. This is one of the more methodologically developed studies in the set, combining perception data with an explicit test-validity lens rather than relying on washback as an unanalysed background assumption.

Onaiba (2014) investigated the washback effect of a revised EFL public examination on teachers' instructional practices, materials, and curriculum use at the University of Leicester, directly testing whether an examination reform altered classroom behaviour — a design that, in principle, allows before/after or perception-based comparison rather than a single cross-sectional snapshot. Abdulhamid (2019) approached a closely related question from the angle of alignment, asking, in a mixed-methods study, what relationship holds between curriculum–examination alignment and washback; this framing usefully reformulates washback not as a fixed property of an examination but as a function of how that examination is aligned with stated curricular aims — a reframing consistent with Orafi and Borg's (2009) earlier finding that Libyan secondary teachers struggled to enact the communicative curriculum precisely since the examination pulled in a different direction. Ali (2024) and Elbakay, et al. (2024) extend the washback strand to specific populations (ninth-grade students and secondary teachers respectively), the latter examining directly how high-stakes examinations shape teaching practice.

Two non-Libyan sources sharpen this discussion by comparison. Rahman, Seraj, Hasan, Namaziandost, and Tilwani (2021) examined the washback of assessment on English teaching–learning practice at Bangladeshi secondary schools and, notably, cite Orafi and Borg's Libyan findings directly as a parallel case of curriculum–examination misalignment producing negative washback — an explicit signal from outside the Libyan literature that the pattern identified in Libya is not distinctive but part of a recognisable class of outcomes in centralised, exam-driven EFL systems. Al-Jaro and Tayeb's (2026) classroom-observation study of how high-stakes tests redefine teaching content adds methodological weight that much of the Libyan literature lacks: direct observation, rather than self-report alone, of how test-driven content substitution occurs in practice. Its inclusion here is deliberate, since it models a design the Libyan literature has yet to widely adopt (see Section 6).

Taken together, the washback cluster is the most internally consistent theme in the literature: every study that addresses it, Libyan or comparative, converges on negative washback and curriculum–examination misalignment as the dominant pattern. Where the studies differ is in how much explanatory weight they place on teachers' own assessment knowledge (a literacy-deficit reading) versus the design and social embedding of the examination itself (a policy/structural reading) — again, Elmahjoub's savviness-versus-tradition question.

4.4 Test Design, Quality, and Curriculum Alignment

A smaller cluster examines the technical quality of Libyan English-language tests themselves, rather than teachers' knowledge or the washback they produce. Masoud (2017) assessed the English-language tests used in Libya's General Certificate of Secondary Education, evaluating aspects of test quality directly relevant to validity and fairness. Saleh and Megaiab (2025) took a design-oriented approach, developing and trialling an English-language test for the secondary certificate with teachers and students in two secondary schools in Al-Bayda city — a rare instance in this literature of research that moves from diagnosis toward a concrete testing intervention rather than stopping at description of the problem. Onaiba's (2024) evaluative study asking whether examination aims and content reflect those of the curriculum returns to the alignment question from the test-design side rather than the washback side, essentially auditing the SECE against curricular intent directly.

This test-design cluster complements the washback cluster usefully: where washback studies infer misalignment from its downstream effects on teaching, this cluster attempts to establish misalignment (or its absence) at the level of the instruments themselves. The two clusters' shared conclusion — that Libyan public English examinations do not adequately reflect the communicative aims of the 2000 curriculum reform — is consequently better evidenced than it would be from either cluster alone, and represents one of the more robust claims a future manuscript drawing on this literature could confidently advance.

4.5 Alternative and Dynamic Assessment

A smaller but forward-looking strand of the literature examines teachers' engagement with assessment approaches that depart from conventional summative testing. Elmahjoub's (2023) study of conceptions of dynamic assessment, already discussed in Section 4.2, belongs here as much as to the beliefs strand. Onaiba and Abudagga (2024) examined the effect of using alternative assessment methods on EFL students' test performance and their attitudes toward those methods, offering outcome-oriented evidence — performance and attitude data — that most of the other studies in this review, which are largely descriptive or perception-based, do not provide. Read alongside the washback and test-design clusters, this strand suggests a literature that is beginning to move from diagnosing the problem (misalignment, negative washback, thin speaking-assessment practice) toward testing potential remedies, though this remains the thinnest and most recent part of the evidence base.

4.6 Higher Education as a Comparative Lens

Omeman (2019) examined teachers' and students' views on the impact of assessment practices on teaching and learning within the Libyan higher-education sector rather than school-level education. Although outside the school-teacher focus of this review's central question, it is included here as a comparative lens: if similar patterns of belief-practice mismatch and examination-driven teaching are visible in Libyan higher education as well as in secondary schools, this would suggest the phenomenon is a feature of Libyan educational assessment culture more broadly rather than one confined to, or explained by, the particular conditions of secondary teacher preparation — a distinction with direct implications for where policy intervention should be targeted.

5. Critical Synthesis: Savviness or Tradition?

Drawing the six thematic strands together, the literature converges strongly on description and diverges sharply on explanation. On description, there is now reasonably convergent evidence, from independent studies using different methods (theses, journal studies, one test-design intervention, one classroom-observation study, and a non-Libyan comparator), that: (a) Libyan school English examinations are not well aligned with the communicative aims of the 2000 curriculum reform; (b) this misalignment produces negative washback, narrowing instruction toward test-preparation activities that can raise examination scores without commensurately improving broader English proficiency; and (c) speaking assessment in particular is thinly and inconsistently practised relative to other skills.

On explanation, the literature has not resolved — and, more importantly, has not always explicitly posed — the question Elmahjoub's review frames most directly: whether these patterns reflect a lack of assessment savviness (that is, a genuine knowledge or literacy deficit among individual teachers, remediable through training) or a policy tradition (that is, a structural feature of a centralised, high-stakes examination system that would produce similar teaching behaviour among knowledgeable and unknowledgeable teachers alike, because the incentives run the same direction regardless of individual literacy). The two readings are not mutually exclusive, and the strongest studies in this set — Ahmed (2019) and Nimehchisalem and Abdalla (2020) in particular — implicitly support a hybrid account, in which teachers may hold reasonably sound assessment beliefs (a literacy asset) that are nonetheless overridden in practice by systemic pressure (a tradition constraint). But few studies test this hybrid account directly, for instance by comparing the classroom practice of teachers with demonstrably higher versus lower measured LAL working under the same examination conditions. Until such a design is attempted for the Libyan school context, the savviness-versus-tradition question will remain, as it is in the current literature, asserted rather than adjudicated.

6. Methodological Critique of the Existing Literature

Several features of the evidence base limit the strength of the conclusions that can currently be drawn. First, the great majority of the primary sources are qualitative case studies or single-institution/single-region theses (Exeter, Sunderland, Leicester, Huddersfield, Carleton), with

small, purposively selected samples; this is entirely appropriate for the exploratory questions these studies ask, but it means the field currently has very little basis for judging how far any given finding — about speaking assessment, about belief-practice gaps, about washback — generalises across Libya's very different regional and institutional conditions (east, west, and south Libya).

Second, most of the evidence rests on self-report — interviews and questionnaires — rather than direct observation of assessment in progress. Nimehchisalem and Abdalla's (2020) own finding that beliefs and practices diverge is itself a warning against relying on self-report alone to characterise practice. Third, the literature is unusually concentrated in published or hard-to-access doctoral theses relative to peer-reviewed journal output; this is understandable given that much of this research has been produced by Libyan doctoral candidates at UK and Canadian universities, but it means the field's cumulative visibility, indexing, and citation is considerably lower than the substance of the work would justify. Fourth, despite the obvious salience of Libya's post-2011 political and institutional instability to any account of teacher training, supervision, and assessment reform capacity, remarkably few of the studies reviewed treat conflict-related disruption as an explicit variable rather than background context. Finally, cross-referencing across this literature has been limited: several washback studies (Ahmed, 2019; Onaiba, 2014; Abdulhamid, 2019; Ali, 2024; Elbakay et al., 2024) appear to have proceeded largely independently of one another despite addressing closely related questions, which is itself part of the rationale for the present synthesis.

7. Gaps and a Future Research Agenda

Four priorities follow from the critical synthesis above. First, the field would benefit from a validated, Libya-specific measure of school teachers' language assessment knowledge — comparable to the Language Assessment Literacy Questionnaire tradition that separates knowledge from belief and from reported practice, allowing the savviness/tradition question in Section 5 to be tested rather than asserted. Second, studies that pair self-report with classroom observation, are needed to establish how far Libyan teachers' assessment knowledge and stated beliefs actually predict observed classroom assessment behaviour. Third, comparative research across Libya's regions would allow the field to distinguish stable structural features of the examination system with very different policy implications. Fourth, building on the small existing intervention strand (Saleh & Megaiab, 2025; Onaiba & Abudagga, 2024), controlled or quasi-

experimental studies that introduce specific LAL training or alternative-assessment interventions and track downstream effects on both teaching practice and student outcomes would move the field from diagnosis toward evidence-based remedy — currently the least developed dimension of this literature.

8. Conclusion

The literature on Libyan school English teachers' language assessment knowledge and practice, though scattered across doctoral theses and regional journals rather than concentrated in a single indexed body of work, is more coherent than its dispersal suggests. Across studies of speaking assessment, assessment beliefs, examination washback, test design, and alternative assessment, a consistent picture emerges of a curriculum reform whose communicative aims were never matched by a corresponding reform of the examination system that governs Libyan secondary English teaching, with predictable consequences for classroom practice. What the field has not yet settled is whether this pattern is best addressed by strengthening individual teachers' assessment literacy or by reforming the structural and policy conditions — the "testing-laden culture" — within which even literate teachers must operate. Resolving that question, through the validated measurement, observational triangulation, and intervention research outlined above, is the most consequential task facing future research in this area.

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.
- Ahmed, A. A. (2019). Washback: Examining English language teaching and learning in Libyan secondary school education (Doctoral dissertation). University of Huddersfield.
- Alderson, J. C., & Wall, D. (1993). Does washback exist? *Applied Linguistics*, 14(2), 115–129.
- Ali, R. (2024). Washback in English education: Insights from Libyan ninth graders. *Journal of English Studies in Arabia Felix*, 3(2), 35–43.
- Al-Jaro, M., & Tayeb, Y. (2026). Beyond the textbook: A classroom observation study of how high-stakes tests washback redefines teaching content. *International Journal of Language Testing*, 16(2).
- Bachman, L. F., & Palmer, A. S. (2010). *Language assessment in practice: Developing language assessments and justifying their use in the real world*. Oxford University Press.
- Barq, A. M. (2026). The difficulties faced by third year university students when learning English online. *Al-Farooq Journal of Sciences*, 2(3), 1661-1667.
- Cheng, L. (2004). The washback effect of a public examination change on teachers' perceptions toward their classroom teaching. In L. Cheng, Y. Watanabe, & A. Curtis (Eds.), *Washback in language testing: Research context and methods* (pp. 147–170). Lawrence Erlbaum Associates.
- Davies, A. (2008). Textbook trends in teaching language testing. *Language Testing*, 25(3), 327–347.
- Elbakay, F. A., Zaglom, K. A., Elkut, K. S., & Farid, A. (2024). The impact of high-stakes examinations on teaching practices: A case study of Libyan secondary school teachers. *African Journal of Advanced Studies in Humanities and Social Sciences*, 505–514.
- Elmahjoub, A. A. (2023). Lack of assessment savviness or policy tradition? Libyan EFL teachers' conceptions of dynamic assessment. *Journal of Academic Research*, 24, 1–9.
- Elmahjoub, A. A. (n.d.). A testing-laden culture: A critical review of English language testing in Libya. [Manuscript / publication details to be confirmed].
- El-Taher Al-Fourganee, Z. (2018). How secondary school Libyan teachers assess students' oral communication (Doctoral dissertation). University of Sunderland.
- Fulcher, G. (2012). Assessment literacy for the language classroom. *Language Assessment Quarterly*, 9(2), 113–132.
- Giraldo, F. (2020). A framework for assessment literacy: Language assessment. *Colombian Applied Linguistics Journal*, 20(2), 179–195.
- Grada, T. K. A. (2014). An investigation into Libyan EFL novice secondary school teachers' current knowledge and practice of speaking assessment: A socio-cultural perspective (Doctoral dissertation). University of Exeter.
- Hasselgreen, A., Carlsen, C., & Helness, H. (2004). European survey of language testing and assessment needs: General findings. CIEP/EALTA.
- Inbar-Lourie, O. (2008). Constructing a language assessment knowledge base: A focus on language assessment courses. *Language Testing*, 25(3), 385–402.
- Kremmel, B., & Harding, L. (2020). Towards a comprehensive, empirically validated framework of language assessment literacy: Developing the Language Assessment Literacy Survey (LALS). *Language Assessment Quarterly*, 17(1), 100–120.
- Masoud, M. M. (2017). The assessment of English language tests in the General Certificate of Secondary Education in Libya. *Al-Mukhtar Journal of Humanities*, 34(1), 1–19.
- Nimehchisalem, V., & Abdalla, M. M. (2020). Assessment beliefs and practices of Libyan secondary school English language teachers. *International Journal of Applied Linguistics & English Literature*, 9(6), 99–109.
- Omeman, A. M. A. (2019). Teachers' and students' views towards the impact of assessment practices on the teaching and learning of EFL in the Libyan higher education sector (Doctoral dissertation).

Language Assessment Knowledge and Practice among Libyan School English Teachers: A Critical Review of the Literature

- Onaiba, A. M. E. M. (2014). Investigating the washback effect of a revised EFL public examination on teachers' instructional practices, materials and curriculum (Doctoral dissertation). University of Leicester.
- Onaiba, A. M. (2024). Do exam aims and content reflect those of the curriculum? An evaluative study. *Asian Journal of Assessment in Teaching and Learning*, 14(1), 55–69.
- Onaiba, A., & Abudagga, Y. F. (2024). The effect of using alternative assessment methods on EFL students' test performance and their attitudes towards these methods. *Journal of Academic Research*, 28(2), 1–10.
- Orafi, S. M. S., & Borg, S. (2009). Intentions and realities in implementing communicative curriculum reform. *System*, 37(2), 243–253.
- 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(1), 12.
- Saleh, H. K., & Megaiab, M. M. (2025). Designing and developing an English language test for secondary certificate in Libya: A case study on Libyan teachers and students in two secondary schools in Al-Bayda City. *North Africa Journal of Scientific Publishing (NAJSP)*, 291–297.
- Taylor, L. (2013). Communicating the theory, practice and principles of language testing to test stakeholders: Some reflections. *Language Testing*, 30(3), 403–412.