

# A BIBLIOMETRIC ANALYSIS OF GLOBAL RESEARCH TRENDS IN ENGLISH LANGUAGE ASSESSMENT AT THE PRIMARY EDUCATION LEVEL (2010–2025)

Somakala A/P Jeryaraman <sup>1\*</sup>  
Izazol Binti Idris <sup>1</sup>  
Nur Fatihah Binti Nordin <sup>1</sup>

<sup>1</sup> Faculty of Human Development, Universiti Pendidikan Sultan Idris, 35900 Tanjong Malim, Perak, Malaysia  
(E-mail: [p20251003406@siswa.upsi.edu.my](mailto:p20251003406@siswa.upsi.edu.my); [izazol@fpm.upsi.edu.my](mailto:izazol@fpm.upsi.edu.my); [p20251003352@siswa.upsi.edu.my](mailto:p20251003352@siswa.upsi.edu.my))

\*Corresponding author: [izazol@fpm.upsi.edu.my](mailto:izazol@fpm.upsi.edu.my)

## Article history

Received date : 19-2-2026  
Revised date : 20-2-2026  
Accepted date : 4-3-2026  
Published date : 14-3-2026

## To cite this document:

Jeryaraman, S., Idris, I., & Nordin, N. F. (2026). A bibliometric analysis of global research trends in English language assessment at the primary education level (2010–2025). *Journal of Islamic, Social, Economics and Development (JISED)*, 11 (81), 255 – 269.

---

**Abstract:** *This bibliometric study analyses global research trends in English language assessment at the primary education level from 2010 to 2025. It is aimed to map the field's growth, intellectual structure, and collaborative patterns. Using data retrieved from the Scopus database, the study identifies the publication trends, influential authors, leading institutions, keyword co-occurrence, co-authorship networks, and country-level research collaboration. The findings from the bibliometric analysis disclose a marked and accelerating increase in research production, especially after 2018, which indicate increased scholarly and policy interest in assessment quality, alignment, and accountability in primary English education. Citation analysis demonstrates that a comparatively small group of highly influential scholars shapes the field in spite of differences in publication volume, meanwhile institutional analysis highlights robust contributions from both Malaysian and internationally renowned universities. Moreover, keyword and network analyses show a conceptually united and internationally collaborative research landscape, dominated by themes such as assessment validity, reliability, formative assessment, assessment literacy, and CEFR alignment. Ultimately, the study provides a systematic overview of the development of English language assessment research in primary education and contributes a strong empirical foundation for future empirical investigations and framework development.*

**Keywords:** *English language assessment; Primary education; Assessment validity*

---

## Introduction

English language assessment in primary education plays a vital role in monitoring students' language development and informing instructional decision-making. At this level, assessment functions not only to measure students' achievement but also to assist learning by identifying strengths, weaknesses, and improvement over time. Effective language assessment, as stressed by Bachman and Palmer (2010), must be aligned with curricular goals and clearly defined language constructs so that it reflects students' communicative competence instead of isolated linguistic knowledge. In primary school settings, where students are still developing basic literacy and cognitive skills, assessments must also be developmentally suitable, fair, and sensitive to students' diverse linguistic backgrounds (McKay, 2006).

In recent years, universal proficiency frameworks such as the Common European Framework of Reference for Languages (CEFR) have gradually influenced English language assessment practices by giving transparent proficiency descriptors and performance standards across language skills (Council of Europe, 2020). The implementation of such frameworks has reinforced the importance on validity, transparency, and comparability in both school-based and large-scale assessments. Nevertheless, research constantly shows that misalignment between assessment tasks, learning outcomes, and proficiency standards can weaken the interpretability and usefulness of assessment results (Fulcher, 2010). Subsequently, there is increasing consensus that English language assessments at the primary level must be thoroughly evaluated to ensure validity, reliability, and alignment with curricular expectations, especially when assessment outcomes are used for high-stakes educational decisions.

In line with that, one of the major issues in primary English language assessment relates to validity and alignment between assessment tasks, curriculum objectives, and the intended language constructs. Past studies have shown that many primary-level assessments place disproportionate emphasis on distinct language elements, such as grammar and vocabulary, instead of communicative competence and integrated language use (Bachman & Palmer, 2010). Eventually, this misalignment might undermine construct validity and may lead to wrong interpretations of students' real language ability. Moreover, assessment tasks that are not developmentally suitable, may disadvantage students in terms of cognitive and literacy limitations instead of lack of language proficiency (McKay, 2006). Subsequently, these issues stress the persistent challenge of developing assessments that are theoretically sound and suitable for primary school students.

In addition to validity concerns, issues related to reliability, fairness, and assessment literacy remain prevalent, especially in contexts where classroom-based or school-level assessments play a central role. Studies show that inconsistencies in test administration, scoring practices, and teacher interpretation of performance standards can deteriorate the reliability of assessment results (Fulcher, 2010). Furthermore, primary students' diverse linguistic and socio-cultural backgrounds increase worries about fairness and accessibility, mainly when assessments depend heavily on reading and writing skills (Shohamy, 2017). Apart from that, limited teacher assessment literacy further aggravates these challenges, as inadequate comprehension of assessment principles may cause teaching-to-the-test practices and irregular implementation across schools (Deluca & Klinger, 2010). Ultimately, these issues highlight the need for systematic evaluation and capacity building to improve the quality and credibility of English language assessment in primary education.

The implications of these challenges are significant for assessment design, classroom practice, and educational policy. The assessments that are not well aligned with curricular goals and language proficiency frameworks may not precisely show students' true communicative ability which eventually leads to misguided instructional decisions and unsuitable placement or progression decisions (Bachman & Palmer, 2010). This, in turn, undermines the credibility and practical value of assessment results for improving teaching and learning. As such, education systems must prioritise the systematic evaluation of assessment instruments to ensure alignment with learning standards, proficiency descriptors, and age-appropriate language use, particularly in summative assessment contexts (Fulcher, 2010).

Similarly, challenges related to reliability, fairness, and teacher assessment literacy have direct consequences for equity and consistency in the classroom. Inconsistent scoring practices and limited understanding of assessment principles among teachers can increase measurement error and lead to uneven implementation that disadvantages certain groups of learners (DeLuca & Klinger, 2010). Inadequate attention to fairness and accessibility may further marginalise students from diverse linguistic and socio-cultural backgrounds (Shohamy, 2017). These implications highlight the importance of sustained professional development in assessment literacy and the establishment of clear, empirically grounded evaluation frameworks to support teachers in designing, administering, and interpreting English language assessments effectively.

In order to address the issues of validity and misalignment, English language assessments in primary education should be thoroughly designed and evaluated using explicit theoretical and proficiency frameworks. As asserted by Bachman and Palmer (2010) assessment tasks must be carefully aligned with curricular expectations, language constructs, and recognised proficiency standards so that results can be interpreted meaningfully. In this regard, frameworks such as the CEFR provide structured performance descriptors that support task design, progression of difficulty, and scoring criteria, thereby enhancing construct validity and transparency [5]. Additionally, consistent evaluation of assessment instruments using document analysis, expert review, and alignment mapping allows in identifying construct underrepresentation and task bias so that the assessments correctly reflect envisioned learning outcomes and students' developmental characteristics (Fulcher, 2010). Collectively, these approaches reinforce the validity and interpretability of primary English language assessments and contribute towards a strong foundation for informed instructional and policy-related decision-making.

Furthermore, concerns related to reliability, fairness, and inconsistent implementation can be mitigated through strengthened teacher assessment literacy and the use of structured evaluation frameworks. This is in line with DeLuca and Klinger (2010) who emphasise that teachers who have strong assessment literacy are able to develop valid tasks, apply reliable scoring practices, and interpret outcomes correctly. Thus, specialised professional development concentrating on rubric use, moderation practices, and fairness principles is crucial, especially in primary education settings where subjective judgement plays an important role. Additionally, the adoption of empirically validated evaluation frameworks can aid teachers and policymakers in evaluating assessment quality systematically, promoting consistency, fairness, and informed decision-making across schools (Shohamy, 2017). Jointly, these solutions reinforce the credibility, usability, and impact of English language assessment in primary education.

Despite increasing attention to issues of validity, reliability, fairness, and alignment in primary English language assessment, existing research remains fragmented and conceptually disconnected, offering limited synthesis of how these concerns have been examined over time

and across contexts. Studies conducted in the past mainly empirical, context-specific, and methodologically varied, thus making it hard to identify main research themes, important theoretical frameworks, collaboration patterns, and underexplored areas in the field (Zupic & Cater, 2015). Furthermore, there is a lack of systematic proof mapping to show which key constructs such as CEFR alignment, cognitive demand, and assessment literacy are represented in primary level English assessment research. Consequently, policymakers and researchers lack a complete overview of research trends, intellectual structures, and gaps essential to enlighten evidence-based framework development.

In response to the fragmented and conceptually dispersed nature of existing research on English language assessment in primary education, a bibliometric analysis is necessary to quantitatively map publication trends, thematic evolution, and research concentrations, thus providing an objective basis for identifying research blind spots and justifying future empirical and framework-based studies (Donthu *et al.*, 2021). Hence, this study employs a bibliometric approach to analyse global research trends in English language assessment at the primary education level from 2010 to 2025, with the aim of providing comprehensive evidence to inform future empirical research and framework development in this field.

In doing so, the following objectives were addressed in this study:

1. To identify and analyse publication trends in English language assessment research at the primary education level from 2010 to 2025, including growth patterns and thematic evolution.
2. To determine the most influential authors in the field by examining publication productivity and citation impact.
3. To identify the leading contributing institutions based on publication output and citation performance in English language assessment research.
4. To examine the intellectual and collaborative structure of the field through:
  - i) keyword co-occurrence analysis,
  - ii) co-authorship analysis, and
  - iii) country-level collaboration network analysis.

## Literature Review

### Language Assessment Literacy as a Foundational Challenge in English Language Assessment

The reviewed literature constantly finds language assessment literacy as a foundational issue determining the quality of English language assessment in educational settings. Weng [20] offers an in-depth synthesis of three decades of language assessment literacy research and proves that teachers' assessment literacy is not one skill but a multi-dimensional construct that include knowledge of assessment purposes, task design, scoring, interpretation, and ethical use of results. Significantly, Weng (2022) further argues that language assessment literacy is very context-dependent, which means that general or imported assessment models may not sufficiently address local classroom realities. This finding is reinforced by Chang (2024) whose empirical study demonstrates that teachers especially in school contexts are found to have significant gaps in technical assessment knowledge and express a need for practical and classroom-based assessment training. Altogether, these studies emphasise a constant structural issue that despite assessment literacy is generally recognised as crucial, it remains unevenly developed and inadequately contextualised for specific educational contexts.

### *Classroom-Level Assessment Practices and the Theory–Practice Gap*

At the classroom level, some studies disclose that assessment practices in primary English education are often pedagogically comprehensive in its purposes but inconsistent in implementation. Mohamed (2021) documents the use of formative assessment strategies such as feedback, questioning, and peer assessment among primary English teachers, yet finds that these practices are not thoroughly linked to summative assessment or broader assessment reformation. Correspondingly, Kholidi (2022) reports that despite teachers effectively adopted digital tools for assessment during and after the pandemic, many struggled to ensure task authenticity, scoring reliability, and consistency, mainly in online environments. In line with that, these findings recommend that innovation in assessment practice is happening, but without strong assessment literacy and structural guidance, such practices is at risk of becoming fragmented and methodologically weak. Thus, the literature shows a gap between assessment policy and its practices in real educational settings.

### **CEFR Alignment Challenges in Policy and Primary-Level Assessment Implementation**

At the policy and curriculum level, research on CEFR implementation discloses more challenges in aligning assessment practices with international proficiency standards. Mohammed (2021) finds a clear mismatch between CEFR policy adoption and classroom readiness by citing problems related to textbooks, teacher preparedness, and contextual adaptation. This misalignment is echoed in Malaysian-based studies by Yunus (2023) and Bayuong and Hashim (2023), which display that despite teachers mostly hold positive attitudes towards CEFR principles, they face real-world constraints that limit effective implementation, mainly in rural and resource-limited contexts. In particular, these studies stress that CEFR descriptors are frequently applied mechanically, without adequate localisation or alignment with students' cognitive and linguistic development at the primary level. This body of work jointly argues that CEFR alignment in primary English assessment remains aspirational instead of fully operational which eventually raises concerns regarding the validity and fairness in the assessment outcomes.

### **Methodological Gaps and the Need for Systematic and Bibliometric Inquiry**

From a methodological and research-design viewpoint, studies have highlighted critical gaps that rationalize further systematic inquiry. Trüb *et al.*, (2025) reveal that rater judgment in primary English writing assessment is greatly influenced by text characteristics, underlining the need for clearer rubrics and structured rater training to enhance reliability. Shafii (2025) further confirms through a scoping review that formative assessment only produces positive learning outcomes when aided by adequate assessment literacy and institutional structures. Meanwhile, systematic and scoping reviews on CEFR implementation in ASEAN contexts (2024–2025) conclude that although CEFR-related research is growing, there is inadequate empirical evidence evaluating how CEFR alignment influence assessment quality, especially in primary education (Hamid *et al.*, 2025). As a whole, these studies disclose a fragmented research landscape dominated by small-scale and context-specific investigations. Indirectly, this fragmentation strengthens the justification for bibliometric and integrative research approaches (Donthu *et al.*, 2021) to map research trends, identify dominant themes and blind spots, and provide a coherent evidence base for developing empirically grounded evaluation frameworks for English language assessment.

## Methodology

This bibliometric study investigates research trends in English language assessment in primary education from 2010 to 2025. According to Zupic and Čater (2015) bibliometrics comprises the systematic organisation and quantitative analysis of bibliographic information retrieved from scientific publications, allowing the identification of publication patterns, influential contributors, and thematic structures within a research domain. Apart from descriptive indicators such as publication year, source journals, and authorship patterns (Donthu *et al.*, 2021), bibliometric techniques also contain advanced analyses such as co-citation and keyword co-occurrence analysis to examine intellectual relationships within the field. In relation to this study, to ensure reliable results, an iterative process of keyword development, database searching, and result refinement was applied (Fahimnia *et al.*, 2015).

Furthermore, to ensure data quality, this study focused completely on peer-reviewed journal articles published in reputable academic journals. According to Mongeon and Paul-Hus (2020), journals indexed in main citation databases are usually recognised as maintaining thorough editorial and review standards, which improve the credibility of published research. Thus, top-tier publications were given priority in this study to support a comprehensive understanding of theoretical developments and research trends in English language assessment. Based on these considerations, Scopus database was chosen as the main data source, as it provides wide coverage of peer-reviewed journals in education, linguistics, and social sciences, together with comprehensive citation and bibliographic data appropriate for bibliometric analysis (Aghaei Chadegani *et al.*, 2013; Donthu *et al.*, 2021).

In addition, to maintain methodological consistency, only journal articles were considered, while books, book chapters, and conference proceedings were omitted so that it is in line with established bibliometric research practices (Zupic & Cater, 2015). Moreover, the search was limited to publications indexed in Scopus from the year 2010 to 2025, permitting the analysis of long-term research trends alongside current developments in the field.

## Data Search Strategy

In doing the bibliometric research, this study employed a systematic screening method to determine appropriate search terms for article retrieval. An early search was conducted in the Scopus database using the TITLE-ABS-KEY field with predefined keywords related to English language assessment. The search terms comprised combinations of “*English language assessment*,” “*language assessment*,” “*English assessment*,” “*ESL assessment*,” and “*EFL assessment*” together with “*primary education*,” “*primary school*,” “*elementary education*,” and related terms. Boolean operators (AND, OR) were applied so that comprehensive coverage of the research domain can be done.

Initially, the search yielded 7,251 records. Then, the results were further refined by applying predefined inclusion and exclusion criteria. Moreover, the search was limited to peer-reviewed journal articles published in English from the year 2010 till 2025, meanwhile review articles, conference papers, book chapters, and other document types were excluded. Finally, after the filtering process only 3,225 articles remained.

Subsequently, the titles, abstracts, and keywords of the chosen records were screened to ensure relevance to English language assessment in primary education. Identical and non-relevant publications were removed during this stage. The final dataset of 2,984, consisting of qualified articles that met all screening requirements, was then used for bibliometric analysis.

### Data Analysis

In doing the bibliometric analysis, BiblioPro: Advanced Bibliometric Analysis System, developed by Dr. Ramlan Mustapha was used. The dataset, exported in CSV formats, was imported into BiblioPro to allow systematic bibliometric mapping and quantitative analysis. A combination of network analysis techniques and descriptive bibliometric indicators was used to address the research objectives and to provide in-depth overview of global research developments in English Language Assessment from year 2010 to 2025.

In order to address the first research objective, descriptive bibliometric analysis was carried out to examine publication development patterns and temporal trends in English language assessment research. Yearly publication output was also analysed to identify periods of increased scholarly activity, meanwhile keyword frequency and thematic evolution analysis were used to explore shifts in research focus over the time. This analysis allowed the identification of main and emerging themes in the field across the specified time span.

Secondly, to answer the second research objective, author-level analysis was performed to determine the most influential authors in English language assessment research. The author efficiency was measured based on the number of publications, meanwhile citation impact indicators were used to recognize highly cited scholars. This dual approach permitted the identification of main contributors who have shaped the development of the field through both research output and scholarly influence.

Thirdly, to fulfil the third research objective, institutional analysis was conducted to find out the top contributing institutions in English language assessment research. The institutions were ranked based on publication productivity and citation performance, which eventually provided information on leading research centres and their relative influence within the international scholarly setting. In doing this analysis, it helps to highlight the institutional patterns of productivity and impact across different regions.

Lastly, in order to address fourth research objective, network analyses were conducted to visualise the intellectual and collaborative structure of English language assessment research. Keyword such as co-occurrence analysis was applied to identify major research themes and conceptual groups within the literature. Moreover, co-authorship analysis mapped collaboration patterns among researchers, meanwhile country-level collaboration networks were analysed to examine international research collaborations and knowledge exchange. Altogether, the integration of the descriptive indicators and network visualisation allowed a structured comprehension of how English language assessment research is organised, linked, and developing at the international level.

### Results

#### Research Trend in English Language Assessment

The Figure 1 shows a clear long-term rising trend in publication production over time, showing a stable growth from a very low baseline to a substantial peak in the most recent year. In the initial years, the number of publications rises gradually, growing from 31 to below 100, with slight fluctuations which indicate a developing but still emerging research area. Halfway through the timeline, growth becomes more visible, reaching around 119 publications, followed by a brief decline (110), which advocates a temporary slowdown instead of a reversal of interest. From this point onward, the trajectory accelerates distinctly, with consistent increase

every single year from 127 to 185, then surpassing 200 and 300 which means heightened scholarly attention and increasing research activity. Although there is a slight decline from 377 to 372, the final data point shows a sharp rise to 506, representing the maximum output in the series. In general, the pattern shows a rapidly growing research field, especially in recent years, signifying growing academic relevance, increased policy or educational interest, and the development of the topic into a well-established area of scholarly inquiry.

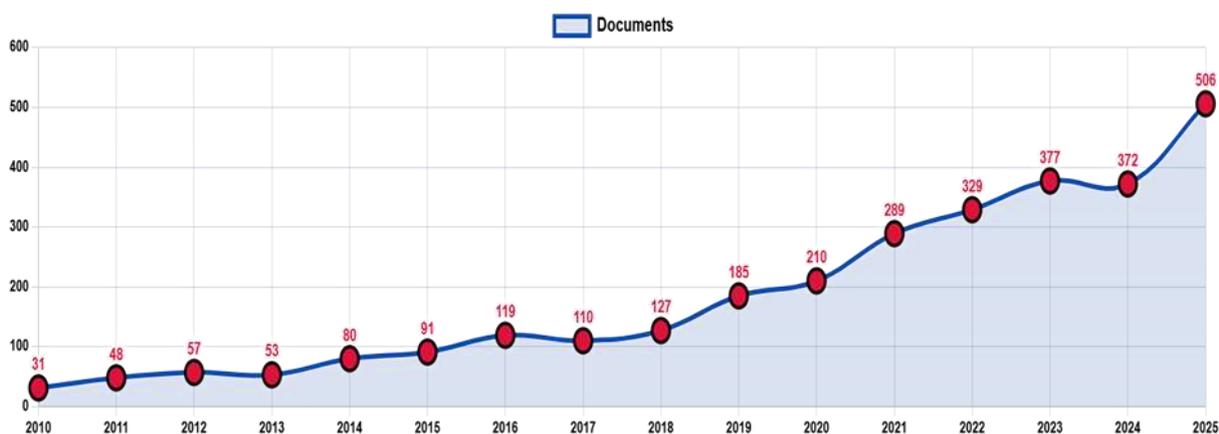


Figure 1: Trend in English Language Assessment

### Most Cited Authors

The table 1 shows the top 10 most productive and influential authors in the field, ranked according to their number of publications, together with total citations and average citations per publication. Verhoeven, L. ranks first with 24 publications and 574 citations, representing continual productivity and robust scholarly effect over time. Despite Jasińska, K., Zhu, X., and Liu, Y. have slightly less publications, their citation counts propose reasonable influence within the field. Particularly, Lee, J. stands out in terms of influence instead of volume, recording the maximum average citations per publication (40.92) and 491 total citations from only 12 publications, signifying that their work is very significant and frequently referenced. Correspondingly, Segers, E. and Tzagari, D. establish strong scholarly impact, with high average citation rates (16.62 and 17.73, respectively), suggesting that their research contributes significantly to theoretical or methodological improvements. Ultimately, the distribution pinpoints that while productivity is vital, research impact in the field is more strongly reflected by citation impact, with some authors employing significant influence in spite of smaller number of publications which indicates high-quality and widely recognised research.

Table 1: Highly Cited Authors in English Language Assessment Research

Rank	Author	Publications	Total Citations	Average Citations per Paper
1	Verhoeven, L.	24	574	23.92
2	Jasińska, K.	15	113	7.53
3	Zhu, X.	14	120	8.57
4	Liu, Y.	13	107	8.23
5	Segers, E.	13	216	16.62
6	Tanoh, F.	12	111	9.25
7	Chen, X.	12	102	8.50
8	Lee, J.	12	491	40.92
9	Liao, X.	11	105	9.55
10	Tzagari, D.	11	195	17.73

### Top Contributing Institutions

The figure 2 demonstrates the distribution of research publications by institutional affiliation, highlighting the contribution of leading universities to the field of English language assessment. Based on the above, Stanford University appears as the most high-volume institution, which records the highest number of documents, indicating robust and continuous research engagement in this area. This is followed closely by Universiti Teknologi MARA, which establishes substantial scholarly output and reflects the prominent role of Malaysian institutions in evolving research on English language assessment. Apart from that, other major contributors consist of Harvard University, Universiti Sains Malaysia, and Universiti Putra Malaysia, all of which show extensive publication volumes, which suggest active research communities and consistent institutional support for assessment-related studies. Furthermore, Oxford University and the National University of Singapore also contribute especially in reinforcing the field's international character. On the other hand, institutions such as Universiti Kebangsaan Malaysia and MIT show comparatively lower publication counts, which indicate more specialised or selective involvement. In general, the pattern discloses a combination of robust regional leadership and global institutional participation, highlighting both the international significance of English language assessment research and the influential role of particular universities in determining scholarly discourse.

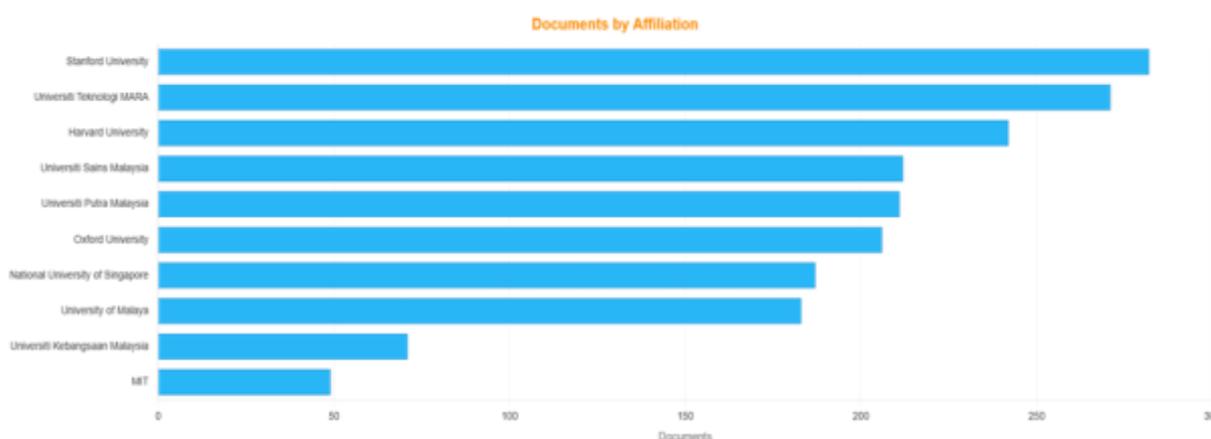


Figure 2: Most Contributing Institutions

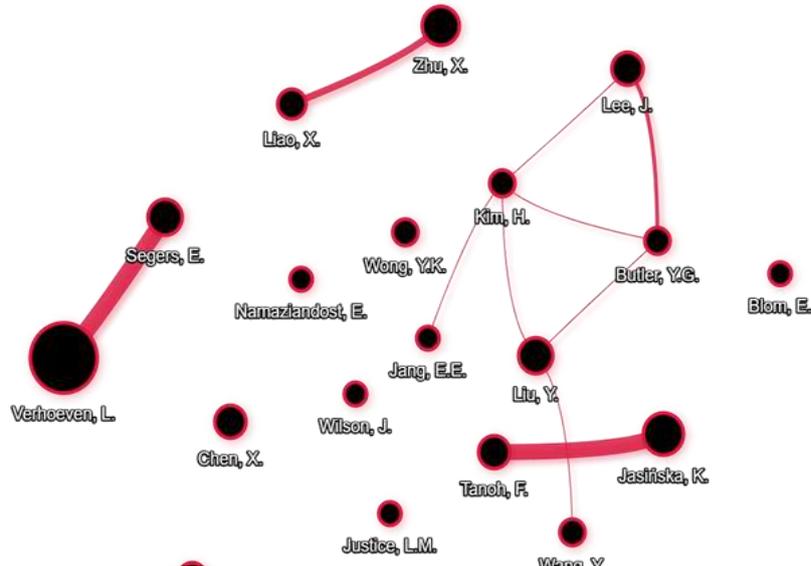
### Key Word Co-Occurrence, Co-Authorship, and Countries' Collaboration

#### Key Word Co-Occurrence in the Field of English Language Assessment

The figure 3 denotes an extremely detailed keyword co-occurrence network that maps the conceptual landscape of research in English language assessment, disclosing how key ideas are interrelated across the literature. Every single node links to a keyword, with bigger nodes representing higher frequency and importance, which means these concepts appear most frequently and play a crucial role in constructing the field. In this study, the most significant central nodes are language assessment, assessment, formative assessment, validity, reliability, assessment literacy, and CEFR. These central nodes function as the main conceptual anchors, which shows that research constantly revolves around issues of assessment quality, interpretation, and alignment with standards. The thick and numerous connecting lines between these nodes show strong co-occurrence, which suggest that studies often integrate multiple assessment components (e.g., validity with formative assessment, or CEFR with reliability) instead of treating them as single constructs. Additionally, the use of different colours emphasizes thematic clusters, where sub-domains such as classroom-based assessment, teacher



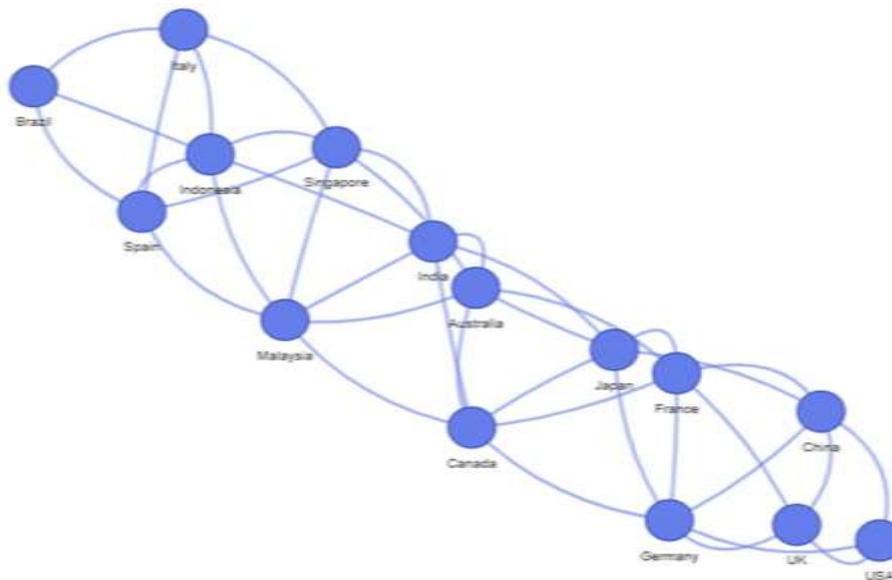
knit collaboration groups play a crucial role in improving knowledge, meanwhile peripheral contributors add depth and thematic diversity to the field.



**Figure 4: Author Co-Authorship Network Based on Bibliometric Analysis of English Language Assessment Studies**

#### Countries Collaboration in the Field of English Language Assessment

The figure 5 depicts a country-level collaboration network that demonstrates international research partnerships in the field of English language assessment. Each node represents a country, meanwhile the connecting lines refer to co-authored publications between researchers from those countries; thicker and more curved relations indicate robust or more regular collaboration. In the above figure, the network forms a continuous, interrelated chain instead of isolated clusters, which indicate that research activity is internationally integrated. Countries like the USA, the UK, Germany, France, Japan, and China inhabit dominant positions and are densely interrelated, emphasising their role as main hubs that anchor international collaboration and enable knowledge exchange across regions. Simultaneously, countries from Asia and the Global South which include Malaysia, Indonesia, Singapore, India, Brazil, and Spain are actively linked to this core, representing growing participation in global research networks instead of peripheral involvement. In addition, the linking position of countries like Australia and Canada indicates their role as mediators linking European, Asian, and American research communities. Collectively, the structure reflects a highly internationalised and collaborative research landscape, where knowledge production in English language assessment is formed by continued cross-national partnerships, with both traditionally leading research countries and emerging countries equally influencing the English language assessment development.



**Figure 5: International Collaboration Network in English Language Assessment Research (2010–2025)**

### Discussion

The bibliometric findings of this study show that research on English language assessment in primary education has gone through significant and continued over the past decade and a half. The sharp rise in publication production, mainly in current years, reflects increasing global concern with assessment validity, alignment with curricular and proficiency frameworks, and the role of assessment in supporting learning instead of just measuring achievement. This trend coincides with extensive educational reforms and the adoption of international frameworks such as the CEFR, which have positioned assessment at the centre of curriculum implementation and accountability. The consistent rise indicates that English language assessment has progressed from a small pedagogical issue into a fundamental research domain in applied linguistics and educational assessment.

In addition, the analysis of authorship and citation patterns shows a centralised intellectual structure, whereby a small number of scholars exert strong influence either through continued productivity or high citation impact. Although some authors were found to contribute widely through a higher number of publications, others demonstrate significant influence in spite of less outputs, which suggest that research quality and conceptual contribution outweigh absolute productivity in determining the field. Moreover, institution-level analysis further strengthens this pattern by showing that both globally dominant universities and Malaysian institutions play a vital role in evolving scholarship. In line with that, the strong presence of Malaysian universities reflects their engagement driven by national assessment reforms and CEFR implementation, meanwhile the contribution of internationally recognised institutions highlights the global relevance of research on primary English language assessment.

Furthermore, the keyword co-occurrence analysis discloses a highly united conceptual landscape, dominated by main constructs such as validity, reliability, formative assessment, assessment literacy, and CEFR alignment. The dense interconnections among these keywords specify that current studies gradually adopt a holistic approach to assessment, including psychometric, pedagogical, and policy-oriented views. At the same time, the relative marginality of keywords linked to fairness, contextual adaptation, and rater judgement

advocates that these areas have been largely overlooked. Correspondingly, co-authorship and country collaboration networks show a strongly internationalised research environment, secured by reputable research hubs (e.g., the USA, UK, Europe, and East Asia) but progressively inclusive of emerging research contexts, especially in Asia. Nevertheless, the focus on influence among a limited group of countries and authors emphasises the need for wider participation and more context-sensitive research perspectives.

### Conclusions

In sum, this bibliometric analysis offers a comprehensive and systematic overview of research developments in English language assessment at the primary education level from 2010 to 2025. The findings suggest that the field has undergone fast expansion, growing conceptual integration, and robust international collaboration. Research focus also has converged around key assessment quality dimensions, predominantly validity, reliability, assessment literacy, and alignment with proficiency frameworks such as the CEFR. Simultaneously, scholarly influence remains focused among a small group of authors, institutions, and countries, which indicate a developed but unevenly distributed research landscape.

Significantly, the study recognizes clear research gaps that permit further investigation, including inadequate empirical focus on contextual adaptation, fairness, cognitive demand, and practical implementation of assessment frameworks in primary classrooms. Through synthesising large-scale bibliometric evidence, this study addresses the fragmentation evident in present empirical research and offers an evidence-based rationale for future studies. The findings also highly support the necessity for empirically validated, contextually grounded evaluation frameworks that can enhance the credibility, usability, and educational impact of English language assessment in primary education. In this sense, this bibliometric analysis not only maps past and present research trends but also provides strategic direction for enhancing theory, practice, and policy in primary English language assessment.

### Acknowledgements

1. The authors would like to express our heartfelt gratitude to the Faculty of Human Development, University Pendidikan Sultan Idris for its academic support and research guidance in completing this study.
2. The authors also would like to acknowledge the Ministry of Education (MOE) for granting the financial support that significantly leads to successful completion of this study.
3. Last but not least, the authors would like to thank Dr. Malar and Dr. Ramlan for their guidance and support throughout completing our research article. Their expertise has not only shaped our research but also allowed us to overcome the challenges faced.

## References

- Aghaei Chadegani, A., Salehi, H., Yunus, M. M., Farhadi, H., Fooladi, M., Farhadi, M., & Ale Ebrahim, N. (2013). A comparison between two main academic literature collections: Web of Science and Scopus databases. *Asian Social Science*, 9(5), 18–26. doi:10.5539/ass.v9n5p18
- 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.
- Bayuonng, P. D., & Hashim, H. (2023). CEFR-aligned primary curriculum framework perceptions. *Malaysian Journal of English Research (MAJER)*.
- Chang, D. Y. S. (2024). Exploring language assessment literacy and assessment needs of English teachers. *Language Testing*.
- Council of Europe. (2020). *Common European framework of reference for languages: Learning, teaching, assessment – Companion volume*. Council of Europe Publishing.
- DeLuca, C., & Klinger, D. A. (2010). Assessment literacy development: Identifying gaps in teacher candidates' learning. *Assessment in Education: Principles, Policy & Practice*, 17(4), 419–438. doi:10.1080/0969594X.2010.516643
- Donthu, N., Kumar, S., Mukherjee, D., Pandey, N., & Lim, W. M. (2021). How to conduct a bibliometric analysis: An overview and guidelines. *Journal of Business Research*, 133, 285–296. doi:10.1016/j.jbusres.2021.04.070
- Fahimnia, B., Sarkis, J., & Davarzani, H. (2015). Green supply chain management: A review and bibliometric analysis. *International Journal of Production Economics*, 162, 101–114. doi:10.1016/j.ijpe.2015.01.003
- Fulcher, G. (2010). *Practical language testing*. Hodder Education.
- Hamid, H. A., Wan Fakhrudin, W. F. W., Afip, L., & Hamid, M. O. (2025). Navigating global standards in local contexts: A systematic review of CEFR-aligned English language curriculum in the ASEAN region. *Journal of Nusantara Studies*, 10(1), 324–352.
- Kholidi, M. A. (2022). EFL teachers' online assessment practices at primary level. Conference paper.
- McKay, P. (2006). *Assessing young language learners*. Cambridge University Press.
- Mohamed, M. (2021). "Assessment for learning" practices amongst primary English teachers. *Pertanika Journal of Social Sciences and Humanities*.
- Mohammed, N. A. (2020). Challenges in implementing CEFR in classrooms. *LSP International Journal*.
- Mongeon, P., & Paul-Hus, A. (2016). The journal coverage of Web of Science and Scopus: A comparative analysis. *Scientometrics*, 106, 213–228. doi:10.1007/s11192-015-1765-6
- Shafii, R. A. (2025). Teacher assessment literacy and formative assessment practices: A scoping review. *Teaching and Teacher Education*.
- Shohamy, E. (2017). Language testing and assessment. In G. Noblit (Ed.), *Encyclopedia of language and education* (2nd ed.). Springer.
- Trüb, R., et al. (2025). Judgment accuracy in primary school EFL writing assessment. *Language Testing*.
- Verbeek, A., Debackere, K., Luwel, M., & Zimmermann, E. (2002). Measuring progress and evolution in science and technology—I: The multiple uses of bibliometric indicators. *International Journal of Management Reviews*, 4(2), 179–211. doi:10.1111/1468-2370.00083
- Weng, F. (2022). Language assessment literacy: A review (1991–2021). *Frontiers in Psychology*.
- Yunus, H. M. (2023). Exploring teachers' beliefs about CEFR-aligned formative assessment in Malaysia. UiTM Institutional Repository.

Zupic, I., & Čater, T. (2015). Bibliometric methods in management and organization. *Organizational Research Methods*, 18(3), 429–472. doi:10.1177/1094428114562629