Trends in English Language Assessment Research: A Bibliometric Analysis (2020-2026)

Birtun BENER[1], Fatma KÖPRÜLÜ[2]

[1]Institute of Graduate Studies, Department of English Language Education, Near East University, Nicosia, Northern Cyprus; Educational Sciences Research Center, Northern Cyprus. *ORCID 0009-0009-8060-574X*

[2]Atatürk Faculty of Education, Department of Educational Administration and Supervision, Near East University, Nicosia, Northern Cyprus; Educational Sciences Research Center, Northern Cyprus. *ORCID 0000-0002-7233-4224*

*Corresponding Author: Birtun BENER

Article History	Abstract
Original Research Article	*This study provides a bibliometric analysis of the research landscape of English language assessment from 2020 to 2025 and articles that were indexed in 2026 to inspect the scope of recently published research articles. Data were extracted from the Scopus database, and bibliometrics were analyzed via VOSviewer. Content analysis was conducted to understand the changes in the research topics. The study reveals the pedagogical, psychological, and technological focus within the English language assessment research landscape. English language assessment research has changed from implementing immediate modifications owing to the COVID-19 pandemic to an AI-focused strategy between 2020 and 2026. Ghiasvand, F. is the leading author, while Khodabakhshzadeh, H., and Maleki, A. are the most cited authors in this domain. Future research must go beyond detecting AI to assessing AI.*
Citation: Birtun BENER, Fatma KÖPRÜLÜ. (2026). Trends in English Language Assessment Research: A Bibliometric Analysis (2020-2026). UKR Journal of Education and Literature (UKRJEL), Volume 2(1), 28-38.	*Keywords:* Language Assessment, EFL, Bibliometric Analysis.

1. Introduction

English language assessment research has shifted from a reactive approach to emergency remote teaching (Mahapatra, 2021; Cheng & Zhang, 2021) due to the COVID-19 pandemic to an AI-integrated framework from 2020 to 2026. In the earlier years, the research focused on survival and addressed teacher technophobia in terms of the struggles they had due to online teaching during the pandemic (Alshakhi & Phan, 2020; Qadir & Omar, 2025). Instructors included gamified online tools (Zuhriyah & Pratolo, 2020) such as Kahoot! to engage students in the online classes (Chiang, 2020). The research has shifted to inclusive and alternative education, which prioritizes language assessment for specialized testing for language learners with special needs (Masitoh & Suryati, 2026) and the replacement of summative assessment with dynamic assessment (Najjarpour & Salimi, 2024) through student-centered approaches such as portfolios (Mahmud et al., 2025) or classroom debates, which promote critical thinking (Lodej, 2020; Eden & Shmila, 2023; Lintangsari & Emaliana, 2020). The scope of the studies has altered from summative approaches to the integration of

Generative AI through the usage of digital tools and ChatGPT as mentors (Zhang & Luo, 2025; Al-Zahrani, 2025; Wiyaka et al., 2024) for learning, rather than considering them as limitations for learning (Ghafouri et al., 2024; Alsagoafi & Alomran, 2025). This alteration necessitated ethical issues of the use of AI (Alavi, 2024) and academic resilience (Cong-Lem, 2026). Therefore, the role of the language instructors has also upgraded from traditional to collaborative assessors with assessment identity (Estaji & Ghiasvand, 2025), in which they started focusing on the process rather than the final product.

This article aimed to identify the main trends, dominant authors, and key theoretical frameworks that define the current state of English language assessment research from 2020 to 2026. The analysis focuses on answering the following questions:

1. What are the main trends in English language assessment research from 2020 to 2026?
2. What are the most commonly used keywords by authors within this domain?

3. What are the main terms used by researchers in this domain?
4. Who are the most productive authors in this domain?

2. Methodology

2.1. Research Design

The bibliometric analysis technique was used for quantitative analysis of published research articles between 2020 and 2025, and also 2026-indexed articles. We obtained quantitative data from the Scopus database. 647 journal articles matched the inclusion criteria. The Scopus data export included publication years, citations, authors' names, countries, keywords, etc. Thematic structures of the selected journal articles were also analyzed. Content analysis was used to analyze articles' abstracts. Therefore, we obtained further understanding on the changes of research topics in English language assessment research.

2.2. Data Selection and Sourcing

As mentioned before, the Scopus database was used to export data on the 8th of December 2025. The search query included three keywords: "language", "assessment", and "EFL". The area was filtered to "Social Sciences" and the type of document to "Journal Articles". The year range was set to 2020 to 2026. Only "English" medium articles were included. At first, 1,088 items were screened. After filtering, 647 articles were selected to be analyzed. The filters on Scopus were the inclusion criteria.

Table 1. Data Retrieval Procedure

Element	Details
Search Query and Filtering (inclusion and exclusion criteria)	TITLE-ABS-KEY (language AND assessment AND efl) AND PUBYEAR > 2019 AND PUBYEAR < 2027 AND (LIMIT-TO (SUBJAREA , "SOCI")) AND (LIMIT-TO (DOCTYPE , "ar")) AND (LIMIT-TO (LANGUAGE , "English"))
Subject Area	Social Sciences
Document Type	Articles
Time Span	2020-2026 (2026 is indexed)
Language	English
Database	Scopus
Documents Screened	1088
Documents included after filtering	647
Date of Data Extraction	8-Dec-25

2.3. Data Analysis

In order to analyze the data, the bibliometric analysis method was used. VOSviewer was implemented for mapping how keywords and terms are interrelated. We analyzed the data by using its network visualizations. A short content analysis was utilized to map the shifts in the topics per year by analyzing the abstracts of several (N=5 approx. per year) research articles found in the dataset.

3. Findings

3.1. Country Productivity
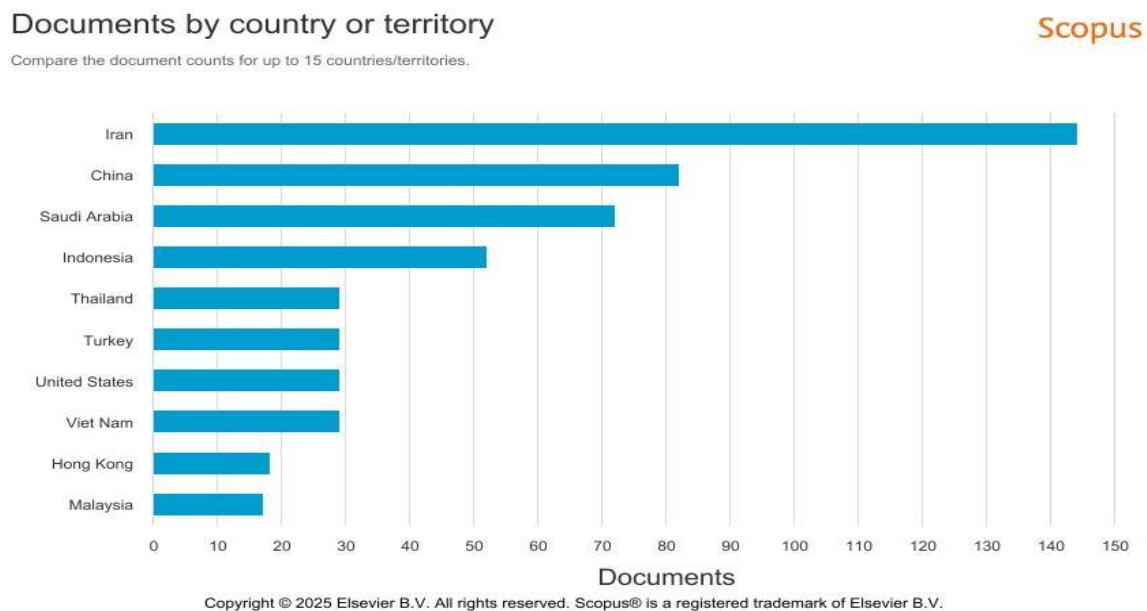
Figure 1. Number of articles published by country

A country-by-country analysis of English language assessment research reveals that Iran leads with N=144 articles, followed by China with over N=80 articles. Saudi Arabia and Indonesia are tied with approximately N=72 and N=52 publications. Thailand, Türkiye, the U.S., and Viet Nam each produced around N=30 articles, while Hong Kong and Malaysia contributed fewer than N=20 each. Overall, this data highlights the significant publication rates of research from Asian and Middle Eastern countries in the English language assessment field.

3.2. Countries' Collaboration Network
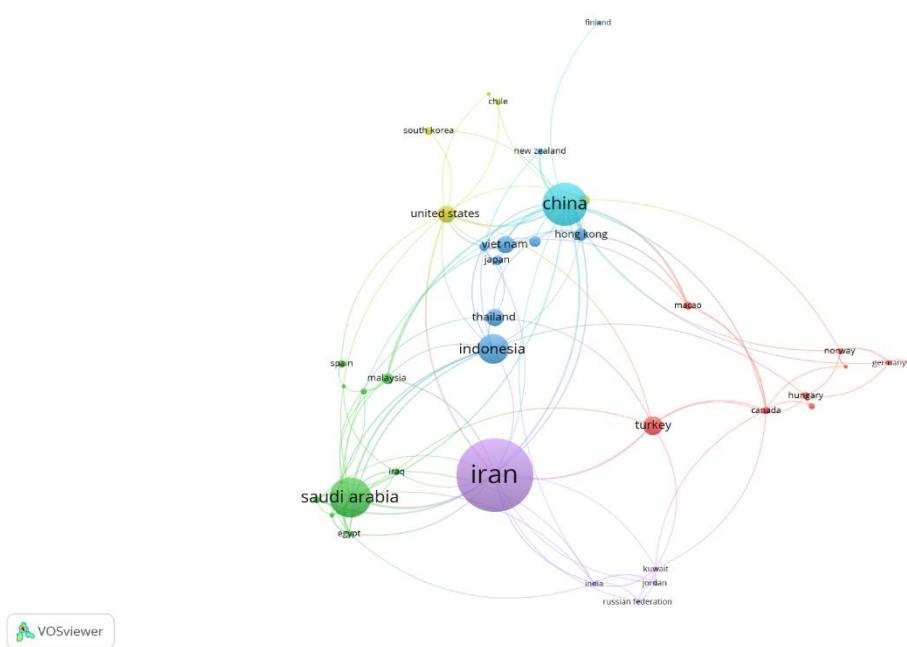
Figure 2. Co-authorship Network

Figure 2 presents the international co-authorship network of publications on English language assessment derived from the Scopus dataset (N= 647). As it is obvious from the figure, Iran and China are the most productive and collaborative countries. Iran has got many connections with Egypt, Iraq, India, Saudi Arabia and Kuwait. That means this present topic is popular among these countries. China is another country that has got many connections, such as Hong Kong, Thailand, Viet Nam,

Japan and the United States. The United States is a smaller contributor; however, it connects many clusters. Türkiye works like a bridge between different countries.

3.3. Geographic Contributions

Researchers from 77 different countries contributed to the English language assessment research. Iran has got the most number of publications with N=144, then China with N=82 publications. Saudi Arabia's research output on this domain is N=72. In the top five countries, Indonesia (617%) and Türkiye (3,80%) come next. The US, the UK, and several European countries contribute less. The US ranks sixth, alongside Viet Nam and Thailand, at 3,44%. The dataset primarily consists of studies from Iran, China, and Saudi Arabia.

3.4. Yearly Publication Output
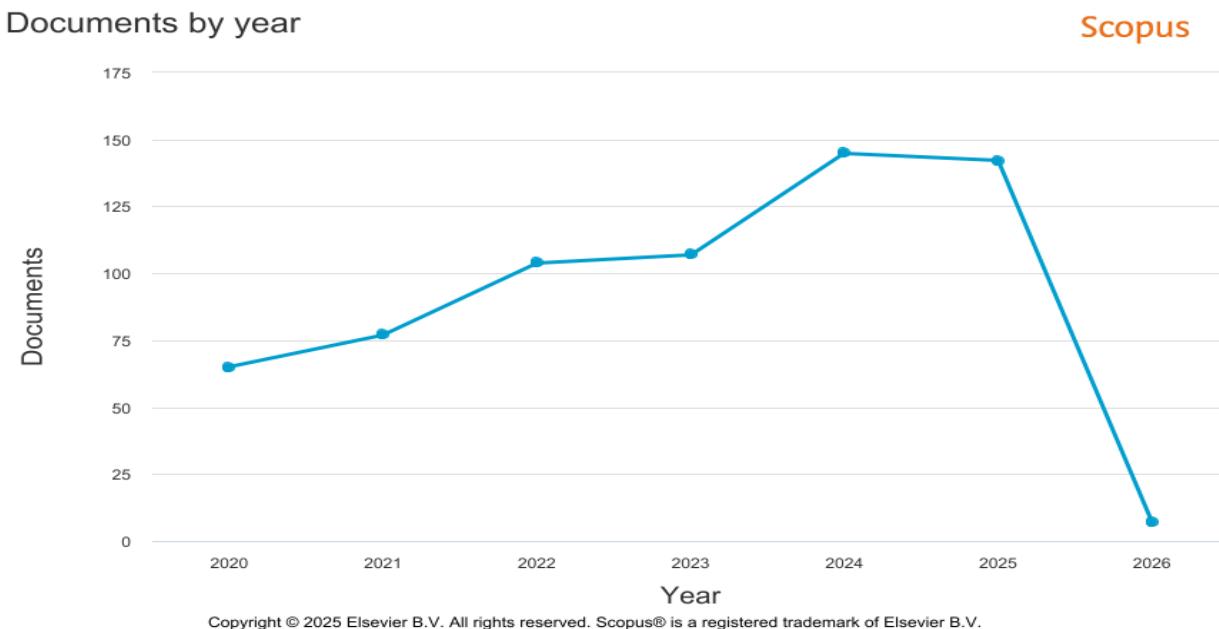
Figure 3. Yearly Publication Output

Figure 3 illustrates that the number of publications on this topic has been steadily growing from 2020 to 2026. Publication quantities have continuously increased, going from N=65 in 2020, to N=104 in 2022, and N=145 in 2024. In 2025, a stable number of research studies were produced (N=142). There are only N=7 indexed articles in 2026 yet. The number of publications has more than doubled between 2020 and 2024, showing that the language assessment research is becoming increasingly important and interesting to researchers.

3.5. Research Question 1: What are the main trends in English language assessment research from 2020 to 2026?
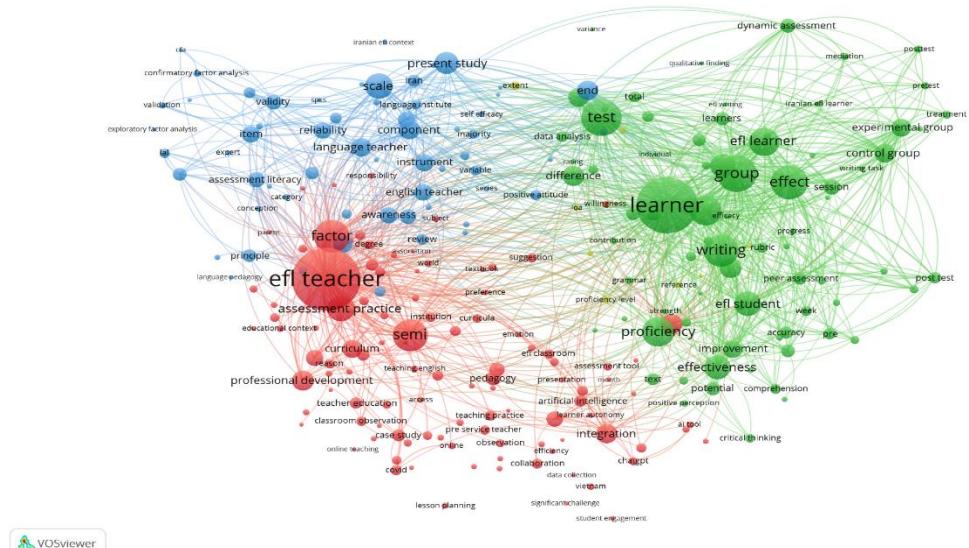
Figure 4. Co-occurrence Network of Titles and Abstracts

Figure 4 presents the co-occurrence network of "titles and abstracts" of N=647 research articles published in Scopus about English language assessment published between 2020 and 2026. The above map shows the size of the nodes, how often a word appears, the connections, and the colors show groupings of themes. The network illustrates the major topics and the thematic relationships in this domain.

The term "learner" (green cluster) is the most common and most connected cluster in the network, making it the most important theme cluster. This group includes writing, proficiency, effectiveness, improvement, peer evaluation, writing rubric, experimental group, control group, pretest, and posttest. Methodological terminology related to experimental and quasi-experimental designs is often utilized in parallel to the term "learner".

Another important cluster that focuses on "EFL teachers" and "language assessment" research is the red cluster. Assessment practice, professional development, teacher education, curriculum, pedagogy, teaching practice, pre-service teachers, and classroom observation are all included in this group. The qualitative and mixed-methods approaches that examine teachers' evaluative beliefs, experiences, and decision-making processes are

highlighted by interviews, case studies, awareness, and perception.

Methodological and psychometric research centered on measurement and instrument development is represented by the third important cluster (blue cluster). This group frequently uses terms like validity, reliability, factor, scale, instrument, item, confirmatory factor analysis, exploratory factor analysis, and assessment literacy. A smaller group of terms (in the upper right green cluster) is focused on dynamic assessment and contains terms like mediation, therapy, pretest, posttest, and qualitative results.

Overall, the co-occurrence network demonstrates that research on language assessment in EFL contexts between 2020 and 2026 is primarily structured around three interrelated thematic domains, which are the learner-centered effectiveness studies, teacher-focused assessment practices and professional development, and finally, measurement-oriented and validation-oriented research. The field is methodologically diverse and adaptable to developing educational contexts, as evidenced by the growing interest in technology-mediated assessment and alternative assessment approaches.

3.6. Research Question 2: What are the most commonly used keywords by authors within this domain?
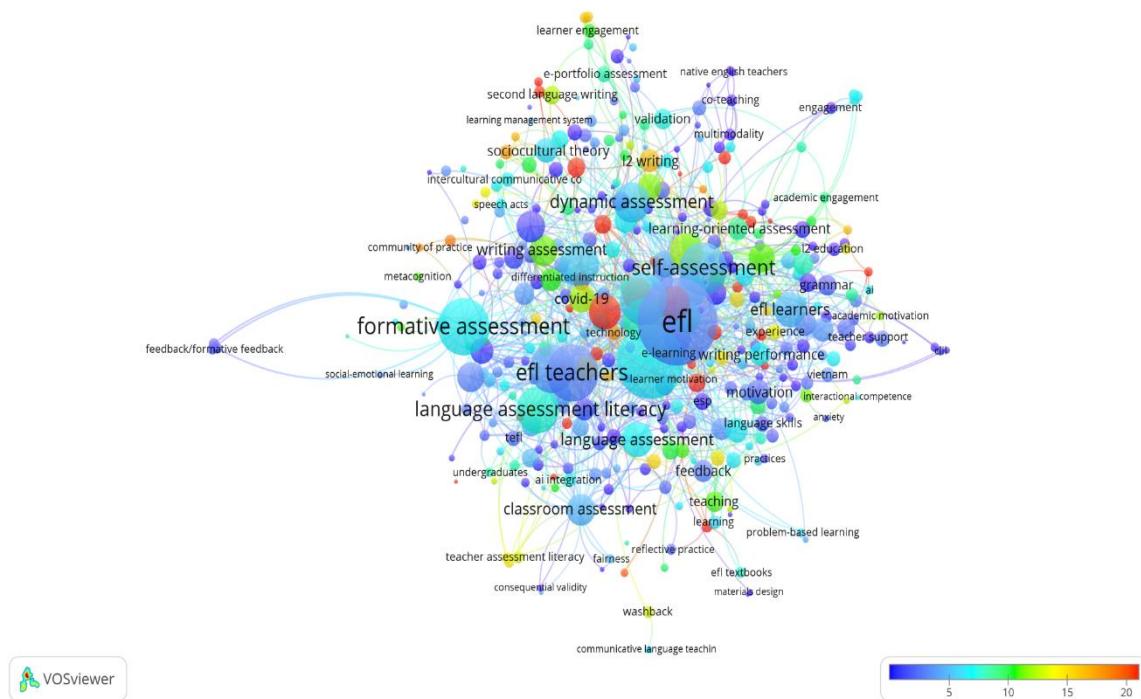
Figure 5. Co-occurrence Network of Keywords

Figure 5 illustrates the co-occurrence network of author keywords in the Scopus-indexed publications on English language assessment between the years 2020 and 2026. The size of the nodes represents how often keywords appear, the links show how often they appear together, and the color shows how often they appear in the same topic. For

instance, "5" (blue) implies that every N=5 article on this topic used the term "EFL" as shown on the color scale on the bottom right of the visualization. The network shows the field's thematic structure.

"EFL" in the middle of the network stands for a vital, connecting node that shows how important it is to connect different disciplines of research. It emphasizes the importance of problems pertaining to EFL teachers in EFL assessment studies, as it relates to assessment literacy, language assessment, and EFL instructors. Terms such as writing performance, grammar, language skills, motivation, and anxiety imply correlations between student outcomes and a learner-centered assessment methodology that influences the cognitive and emotional dimensions of language acquisition. One of the most important groups of themes (turquoise) includes formative and learning-oriented assessment. Formative assessment, self-assessment, feedback, classroom assessment, and learning-oriented assessment are some terms that are very similar to each other.

Another big group focuses on different and changing ways to assess, with "dynamic assessment" (light blue) becoming a well-known term linked to sociocultural theory, writing assessment, L2 writing, and mediation.

The network also shows that a new group is forming around technology-enhanced testing (orange). E-learning, e-portfolio assessment, technology, AI, and AI integration are some of the words that this group uses. These words are connected to both teacher- and learner-focused keywords. This shows that people are becoming more interested in how AI and digital technologies are being used in tests. The emergence of COVID-19 within this cluster further exemplifies the impact of recent global disruptions on assessment methodologies and research priorities.

To sum up, the keyword co-occurrence network shows that language assessment research in EFL contexts from 2020 to 2026 is conceptually organized around formative and learning-oriented assessment, teacher assessment literacy, learner engagement and motivation, dynamic assessment, and the growing role of technology in assessment practices.

3.7. Research Question 3: What are the main terms used by researchers in this domain?
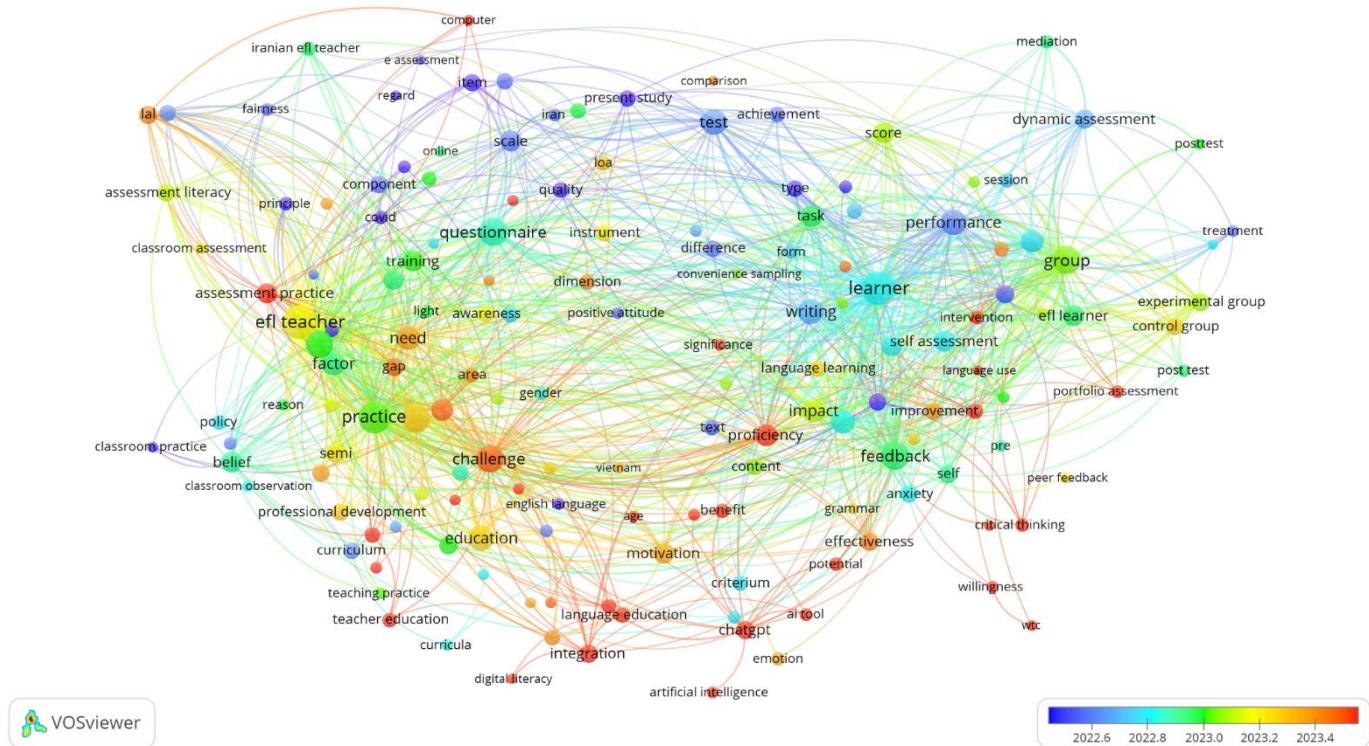
Figure 6. Co-occurrence Network of Terms

The co-occurrence analysis network of terms is illustrated on Figure 6. The network shows the importance put on "teaching practice" and "teacher competence" by the researchers. "Formative assessment" and "peer assessment" are the most common, typically with "feedback" and "self-assessment" as well. A big group is also focused on "language assessment literacy" and "teacher education," which shows how important it is for EFL teachers to know how to use or create tests. The network also has a quickly growing technological side, where "ChatGPT" and "Artificial Intelligence" are connecting together. Terms such as "questionnaire", "factor analysis", and "post-test" imply the research articles were about researchers' development of scales and their scales' validation. To sum up, this figure shows that this topic is investigated in numerous aspects.

3.8. Research Question 4: Who are the most productive authors in this domain?

3.8.1 Leading Authors
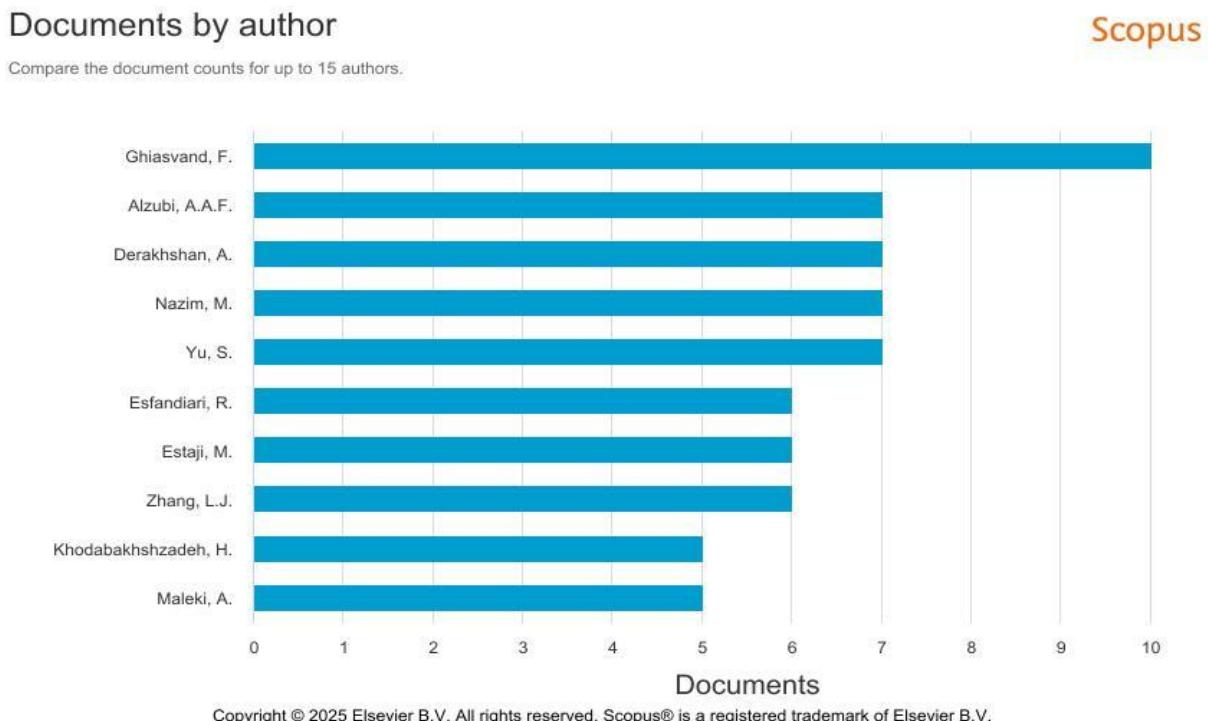
Figure 7. Number of articles published

Figure 8 is a bar chart generated by Scopus titled "Documents by author". It shows how productive researchers are in the dataset we used for our bibliometric analysis. The writers' names are shown vertically, and the number of papers (publications) linked to each author, from 0 to 11, is listed horizontally. The bibliometric analysis of author productivity by using Scopus data finds the 10 most productive authors in the discipline. F. Ghiasvand is the top researcher with the most work, having written 10 research articles. Next, there is a separate group of four researchers (Alzubi, A.A.F., Derakhshan, A., Nazim, M., and Yu, S.) who all have 7 publications and share the second place. Esfandiari, R., Estaji, M., and Zhang, L.J. each wrote six papers, while Khodabakhshzadeh, H., and Maleki, A. wrote five papers each. In general, Ghiasvand, F. has a clear lead in the number of publications. However, the fact that the other top authors have similar results suggests that there is a very active group of researchers.

#### 3.8.2 Bibliographic Coupling Network of Authors
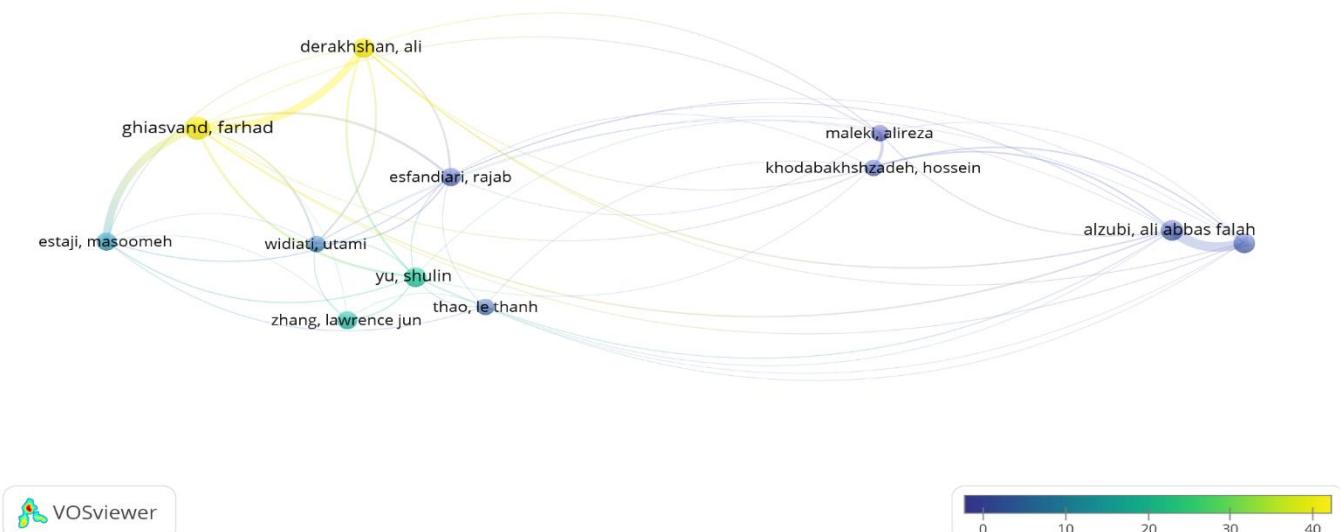
Figure 8. Bibliographic Coupling Network of Authors

A bibliographic coupling analysis was conducted to identify the key authors and the intellectual communities they form based on shared theoretical foundations. The network shows three main groups of important researchers. The first group, which includes Derakhshan, A., Ghiasvand, F., and Estaji, M., is in charge of most of the study on the psychological and emotional aspects of English language assessment. This group of authors showed a strong relationship with over approximately N=35 shared references. A second, distinct cluster is formed by Yu, S., and Zheng, Y., who focus primarily on writing assessment and feedback literacy. Their considerable bibliographic connection underscores a common theoretical basis in social-constructivist theories of feedback. Finally, Alzubi, A.A.F., and Nazim, M. stand out as a unique, very linked pair with the strongest connection in the network (N=68 shared references). This shows that they are working on a focused area of study with almost the same theoretical approach.

3.8.3 Co-citation Network of Authors
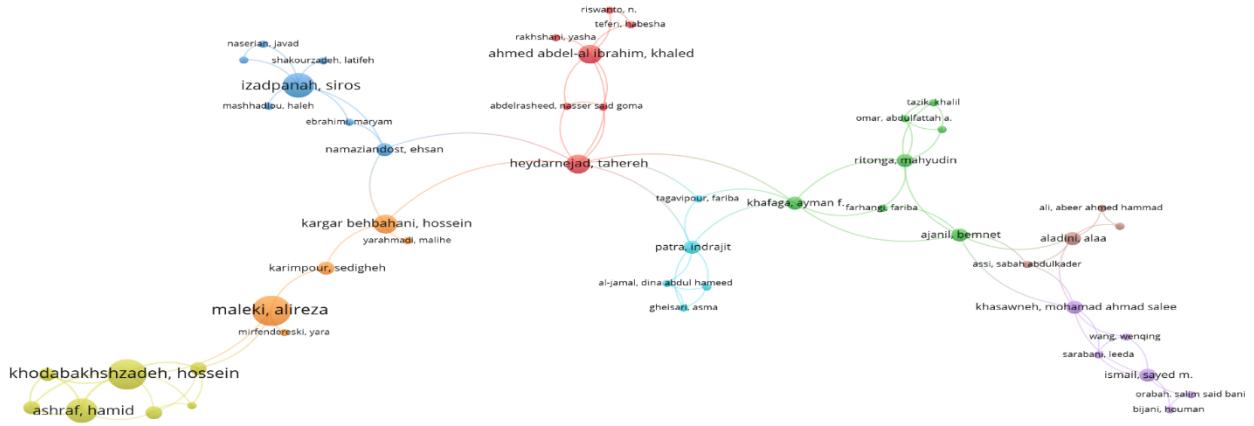
Figure 9. Co-citation Network of Authors

The co-citation network shows that *Khodabakhshzadeh, H.* and *Maleki, A.* anchor the yellow/orange research theme, while *Heydarnejad, T.* acts as the central point for the entire network, connecting the different clusters of authors. *Khodabakhshzadeh, H.* and *Ashraf, H.* anchor the yellow cluster, representing the most cited foundation of the map, with the strongest co-citation link between them. This group connects to *Maleki, A.* in the orange cluster, who serves as a major citation hub alongside *Kargar Behbahani, H.*. Acting as the central pivot for the entire network, *Heydarnejad, T.* in the red cluster, bridges these authors to *Ahmed Abdel-al Ibrahim, K.* and the broader green cluster, where *Ritonga, M.* and *Ajanil, B.* are prominent. Meanwhile, the blue cluster is clearly defined by the influence of *Izadpanah, S.*, who is frequently cited alongside *Mashhadlou, H.*, *Ebrahim, M.*, *Naserian, J.* and *Shakourzadeh L.*.

4. Discussion

It can be said that the number of publications is rising, where it reached to the highest point between 2024 and 2025. Iran (17.08%), China (9.73%), and Saudi Arabia (8.54%) bring the most number of publications to the English language assessment research. These percentages can explain the pedagogical implications of these countries in the field. "EFL" and "Language assessment literacy" keywords have the highest frequency. The main thematic structure of the English language assessment research changed to on-going process assessment instead of summative assessment. AI-integrated assessment literacy is also an important term among researchers (Reyhaneh & Behzad, 2025). That is, they are more interested in understanding what language instructors assess and how well they know how to assess. Usage of "self-assessment", "peer-assessment", and "dynamic assessment" keywords makes it obvious that there is a current change from teacher-centered assessment to shared accountability (Al-Obaydi & Pikhart, 2025; Csöregh & Erickson, 2025). The relationship between the keywords "EFL teachers" and "professional development" implies that assessment is now more important in teacher education (Shokpour et al., 2024). Assessment should not be seen as an administrative duty. The co-authorship network shows that Iranian and Chinese researchers are building collaboration clusters, demonstrating that while ideas are imported, application and empirical data remain restricted. These researchers are cultivating a particular knowledge base that addresses cultural challenges such as substantial class sizes and intense examination pressures in Asia.

4.1. The Evolution of Trends in English Language Assessment Research

From 2020 to 2026, traditional language assessment approaches had to be changed due to the COVID-19 pandemic. Researchers started to implement interactive digital tools in their online classes (Chiang, 2020; Zuhriyah & Pratolo, 2020), and there was an increasing focus on assessment literacy from 2020 to 2021, when emergency remote teaching was a big problem (Mahapatra, 2021; Cheng & Zhang, 2021). Some of the main worries were teachers' "technophobia" and "emotions" (Qadir & Omar, 2025; Alshakhi & Phan, 2020; Tajabadi & Meihami, 2024). In the years after the pandemic (2022–2023), research became more inclusive by focusing on students with special needs (Lodej, 2020; Eden & Shmila, 2023; Lintangsari & Emaliana, 2020) and switching from summative to formative evaluation techniques (Yung, 2020; Tsao, 2025). By 2024, AI, including ChatGPT and other tools (Ghafouri et al., 2024; Alsagoafi & Alomran, 2025), was at the center of attention in terms of the effective use of AI for teaching and assessment (Topuz et al, 2025; Wiyaka et al., 2024). Other topics they were also interested in were "academic resilience" and "emotional labor" (Khasawned et al., 2025). Nurisma (2025) explored the perceptions of experienced and novice language instructors on online English reading skills classes. And finally, online cheating is now the hot topic for researchers (Cong-Lem, 2026; Maleki et al., 2026) and also language instructors to have an assessment identity (Estaji & Ghiasvand, 2025; Rezai, 2024). Lu et al. (2026) put emphasis on teacher and student collaboration for the assessment in their recent research article. Mahmud et al. (2025) studied the undergraduate EFL students' higher order skills through portfolio assessment in writing skills lessons. Another study discovered the relationship between EFL learners' classroom belonging, emotion regulation, and their perceptions of online English classes (Yüksel et al., 2025).

5. Conclusion

There is an important diversification in many aspects of English language assessment research. Iran, China, and Saudi Arabia are contributing to the field in terms of how learning is affected by assessment. Thematic landscape can be explained by a huge focus on teachers' assessment literacy. The researchers' main goal is to grant the language instructors the skill to use formative procedures such as peer feedback and self-assessment. They believe that formative assessment can help students become more autonomous (Noorwali, 2025) and learn new languages easily. Instructors and their methods of assessment are the main participants of the research conducted between 2020 and 2026. According to Shao (2025) and Yang & May (2023), university instructors' and pre-service language teachers' assessment literacy and professional development are more important than students' achievement. English language assessment research is mainly conducted with qualitative research methods, including semi-structured interviews and thematic analysis. As mentioned before, the main subject of research is teachers' assessment literacy. Some studies also implemented quantitative research methods in order to validate the scale they developed. They mainly conducted structural equation modelling analysis (SEM) on their newly developed instruments, and then they used that scale to inspect the relationship between language instructors' assessment literacy and their pedagogical practices, and also AI-integrated assessment (Hastomo et al., 2024; Zheng et al., 2026; Alshakhi, 2024).

References

1. Alavi, S. (2024). Investigating the Impact of Ethical Considerations in Final Assignment Submission among Postgraduate EFL Students. *Journal of Multidisciplinary in Social Sciences*, *20*(1).

2. Alshakhi, A. (2024). Speaking Skill Assessment Instrument Validity: An Investigation into Instructors' Perceptions. *Journal of Ecohumanism*, *3*(7), 4203-4207.

3. Alsagoafi, A. A., & Alomran, H. S. (2025). Revolutionizing Assessment: Leveraging ChatGPT for Automated Item Generation: An AI Driven Exploratory Study with EFL Teachers. *World Journal of English Language*, *15*(6), 385-385.

4. Al-Obaydi, L. H., & Pikhart, M. (2025). AI partner versus human partner: comparing AI-based peer assessment with human-generated peer assessment in examining writing skills. *Language Testing in Asia*, *15*(1), 38.

5. Al-Zahrani, A. A. (2025). Exploring the Role of ChatGPT in English Teaching and Learning: Educators' and Language Learners' Perspectives. *Educational Process: International Journal*, *17*, e2025330.

6. Chiang, H. H. (2020). Kahoot! in an EFL reading class. *Journal of Language Teaching and Research*, *11*(1), 33-44.

7. Cong-Lem, N. (2026). "Is this Really your Work?": A Qualitative Study of Teacher-Led Interviews and Student Accountability in the Age of Generative AI. *Journal of Academic Ethics*, *24*(1), 29.

8. Csöregh, A. M., & Erickson, G. (2025). Fairness and Accountability: A model of language assessment identity positions Rättvisa och ansvarighet. *Acta Didactica Norden*, *19*(1).

9. Estaji, M., & Ghiasvand, F. (2025). Beyond the gradebook: embracing the potentials of teacher assessment identity (TAI) in (re) shaping English language professors' professional development and success in higher education. *Language Testing in Asia*, *15*(1), 48.

10. Ghafouri, M., Hassaskhah, J., & Mahdavi-Zafarghandi, A. (2024). From virtual assistant to writing mentor: Exploring the impact of a ChatGPT-based writing instruction protocol on EFL teachers' self-efficacy and learners' writing skill. *Language Teaching Research*, 13621688241239764.

11. Hastomo, T., Mandasari, B., & Widiati, U. (2024). Scrutinizing Indonesian pre-service teachers' technological knowledge in utilizing AI-powered tools. *Journal of education and learning (EduLearn)*, *18*(4), 1572-1581.

12. Mahmud, T., Rabbi, M. F., Hossain, M. T., Talukder, A. A., & Hasan, M. K. (2025). Portfolio assessment for developing higher order thinking skills in Bangladeshi undergraduate EFL writing classes. *Language Testing in Asia*, *15*(1), 1-18.

13. Maleki, A., Khodabakhshzadeh, H., Zeraatpishe, M., & Baghaei, P. (2025). Development and validation of the online formative assessment validity scale (OFAVS) in EFL contexts. *Discover Sustainability*, *6*(1), 471.

14. Masitoh, F., & Suryati, N. (2026). Attending Inclusive EFL Teachers Identity for Students with Special Educational Needs. *Language Related Research*, *16*(3), 119-145.

15. Najjarpour, M., & Salimi, E. A. (2024). Navigating mindset trajectories: Exploring EFL teachers' evolution in embracing dynamic and summative assessment in the language classroom. *International Journal of Language Testing*, *14*(2), 117-132.

16. Noorwali, S. (2025). Investigating Technology Use and Learners' Autonomy among Saudi EFL Learners.

17. Nurisma, R. A., Dewi, O. T. S., Irawati, R., Sari, R., Suryati, N., & Widiati, U. (2025). Language Assessment Literacy in Online EFL Reading Class: What Novice and Experienced Teachers Reveal. *Computer-Assisted Language Learning Electronic Journal*, *26*(1), 129-152.

18. Phuong, H. Y., Tran, N. B. C., Nguyen, T. T. L., Lam, T. D., Bui, N. Q., & Le, T. T. (2025). From tests to tasks: how Vietnamese EFL teachers navigate washback through formative assessment practices. *Language Testing in Asia*, *15*(1), 45.

19. Reyhaneh, B., & Behzad, G. (2025). Examining the e-assessment literacy of Iranian English as a foreign language teachers. *Journal of Research in Curriculum Instruction and Educational Technology*, *11*(1), 379-307.

20. Qadir, S. M., & Omar, R. M. (2025). Digital literacy, technophobia, and technophilia: Exploring their roles in artificial intelligence integration into English language education. *E-Learning and Digital Media*, 20427530251395634.

21. Rezai, A. (2024). The role of teacher assessment literacy in job stress and job burnout in EFL contexts: a mixed-methods investigation. *Asian-Pacific Journal of Second and Foreign Language Education*, *9*(1), 3.

22. Shao, S. (2025). Developing language assessment literacy: Formative assessment practice from pre-service teacher to novice teacher. *The Asia-Pacific Education Researcher*, *34*(4), 1507-1517.

23. Shokpour, N., Eslampanah, N., Asadi, V., O'NEILL, S. H. I. R. L. E. Y., Dann, C., & Wang, J. (2024). Iranian English as a Foreign Language (EFL) teachers' professional development needs in educational contexts: Exploring the challenge of reflective practice. *Journal of Advances in Medical Education & Professionalism*, *12*(4), 261.

24. Tajabadi, H., & Meihami, H. (2024). "The More, the Merrier; the More Ideas, the Better Feeling": Examining the Role of Creativity in Regulating Emotions among EFL Teachers. *Open Education Studies*, *6*(1), 20240007.

25. Topuz, A. C., Yıldız, M., Taşlıbeyaz, E., Polat, H., & Kurşun, E. (2025). Is generative AI ready to replace human raters in scoring EFL writing? Comparison of human and automated essay evaluation. *Educational Technology & Society*, *28*(3), 36-50.

26. Wang, Z. (2025). Redefining Language Education: How Integrating EFL Teacher-Student Relationships With AI-Powered Classrooms Impacts Emotional Health and EFL Learners' Depression. *European Journal of Education*, *60*(4), e70347.

27. Wiyaka, W., Silitonga, L. M., Sunardi, S., & Pramudi, Y. T. C. (2024). From Nervous to Fluent: The Impact of AI Chatbot-Assisted Assessment on English Reading Anxiety and Performance in Indonesia. *Theory and Practice in Language Studies*, *14*(12), 3851-3860.

28. Yang, R., & May, L. (2023). Understanding and implementing learning-oriented assessment: Evidence From three Chinese junior high school

EFL teachers. *Chinese Journal of Applied Linguistics*, *46*(2), 308-325.

29. Yüksel, H. G., Solhi, M., Özcan, E., & Giritlioğlu, N. B. (2025). The associations between EFL learners' L2 class belongingness, emotion regulation strategies, and perceived L2 proficiency in an online learning context. *The Language Learning Journal*, *53*(3), 306-322.

30. Zhang, L., & Luo, T. (2025). Guided or guiding: contradictions and conflicts in AI-Assisted second language writing for EFL learners from the activity theory perspective. *Innovation in Language Learning and Teaching*, 1-22.

31. Zheng, Y., Liu, C., & Yu, S. (2026). "How should I assess their writing? It'sa headache for me.": Understanding teacher assessment literacy in the collaborative writing context. *Teaching and Teacher Education*, *171*, 105343.

32. Zuhriyah, S., & Pratolo, B. W. (2020). Exploring Students' Views in the Use of Quizizz as an Assessment Tool in English as a Foreign Language (EFL) Class. *Universal Journal of Educational Research*, *8*(11), 5312-5317.