

ASSESSMENT LITERACY IN LANGUAGE EDUCATION: THE STORY SO FAR

Nguyen Thi To Loan^{1,*}, Duong Thu Mai²

¹*Faculty of Foreign Languages, Hung Vuong University, Phu Tho, Vietnam*

²*Faculty of English Language Teacher Education, VNU University of Languages and International Studies,
Pham Van Dong, Cau Giay, Hanoi, Vietnam*

Received 22 May 2023

Revised 10 August 2023; Accepted 26 August 2023

Abstract: Over the past three decades, there has been a considerable increase in research of assessment literacy in language education or language assessment literacy (henceforth LAL), as evidenced by academic researchers' growing interest and publications. Given the importance of LAL in the language teachers' professional development, this paper adopts a systematic review protocol to present a comprehensive literature review on LAL. Based on the analysis of 211 relevant publications identified through the Scopus search engine and using VOSviewer software for bibliometric analysis, this review sheds light on perspectives that have not been thoroughly explored or evaluated by previous studies on the subject. The findings have several implications for the current state of LAL literature, indicating potential future research directions and highlighting existing research gaps. These results provide a robust framework for a deeper understanding of the evolution of research topics, scientific methodologies, and trends in this cutting-edge and captivating field of study.

Keywords: language education, assessment literacy, bibliometric analysis, systematic review

1. Introduction

In the classroom, teachers often spend between one-third and fifty percent of their time on assessment (Cheng et al., 2004) such as planning, creating, carrying out, evaluating assessment and “use assessment evidence to inform their teaching” (Black & Wiliam, 2018, p. 553). They are doing a most significant yet most challenging job, and to be qualified for this job, they need to have sufficient assessment literacy, or language assessment literacy, specifically for language education (Inbar-Lourie, 2013). Assessment literacy in language education or Language Assessment Literacy (LAL) is the term used to describe the knowledge, abilities, and values that stakeholders need to have in order to manage a variety of language assessment tasks (Taylor, 2013). It can be a knowledge base (Malone, 2013), an ability (Deygers & Malone, 2019), or a competence (Baker & Riches, 2018). Teachers who possess such skills, prowess, or knowledge are typically seen as being literate in language assessment. They are able to conduct valid and reliable assessments, maximize the use of evaluations to guide teaching and learning, and make wise decisions on the language proficiency of their students. In other words, LAL among teachers can act as a driving force for effective language instruction. Although they are most in need of it, language teachers are not the only ones who experience LAL (Harding & Kremmel, 2016). In recent years, language testing and evaluation have gained a more prominent position in educational, political, business, and social agendas

* Corresponding author.

Email address: nguyenthitolan@hvu.edu.vn

(Inbar-Lourie, 2017). The significance of teachers' LAL has captured the interest of scholars in recent years, as evidenced by works such as Taylor (2013) and Vogt & Tsagari (2014). Also, research on teachers' LAL has produced valuable insights with substantial implications for language teacher education and professional development.

Despite the conceptual importance of LAL in language testing and assessment, there is a lack of a comprehensive picture in the literature that outlines the extent of LAL research, the aspects of LAL that have been or have not yet been covered, the methods used to investigate these aspects, and the contributions LAL research has made to language teacher education and professional development. Moreover, within the existing conceptual discussions surrounding LAL, such as those put forth by scholars like Harding and Kremmel (2016) and Tsagari (2020), some arguments related to LAL research, such as its geographical sources and methodological designs, appear to lack empirical evidence.

2. Literature Review

First, the scholarly interest in "assessment literacy" can be traced back to the implementation of test-driven accountability systems in various educational contexts during the 1980s. Within these educational settings, educators were obligated to monitor and report the achievement of all students in accordance with established criteria. Following a comprehensive examination of informal language assessment methods, Brindley (2001) introduced the concept of LAL to specifically address the distinct attributes and needs associated with language-related disciplines.

Subsequent LAL research has focused on conceptual frameworks and relationships among examined elements. Davies (2008) and Inbar-Lourie (2008) established an early framework with principles, knowledge, and skills. Taylor (2013) devised an LAL framework for diverse stakeholders, including eight elements. Baker and Riches (2018) refined it to seven elements with Haitian teachers. As empirical data continues to accumulate and undergo validation in various assessment settings, new conceptual frameworks are constantly being constructed (Yan & Fan, 2021).

Stakeholder groups in language testing and assessment have varying interests, needs, and expectations, resulting in differences in their grasp of LAL frameworks and their proficiency in specific components (Benjamin Kremmel & Luke Harding, 2020). These differences have led to a tendency not to study LAL as a comprehensive concept but rather as separate investigations within distinct stakeholder groups. Among these groups, classroom teachers have been the most extensively examined in the academic literature. Stakeholders' LAL is shaped by a combination of internal and external factors. Internal factors, such as teachers' self-confidence, willingness to participate in assessment training, and teaching experience, directly influence LAL. External factors, encompassing educational environments, administrative orders for teaching and assessment, educational policies, and socio-cultural values in language teaching, also significantly impact LAL.

On Web of Science (WoS) and Scopus, three LAL review papers were discovered. Through an analysis of 100 research from 1985 to 2015, Xu & Brown (2016) revised how teachers measure literacy in general education. They put up a framework for how teachers might effectively use assessment literacy, which they connected to teacher professional development. To examine a broad trend in LAL research, Lee & Butler (2020) synthesized 52 empirical works on the topic. They discovered that questionnaires and interviews were the most common techniques of data gathering, and that the majority of LAL studies targeted in-service teachers. As a result, they recommended taking into account the viewpoints of learners and incorporating various study techniques to better

comprehend the nature of LAL. Similar to this, Coombe et al. (2020) concentrated on assessment training and assessment literacy for teachers. They learned from the review that LAL had many facets and that assessment training ought to be a crucial component of teachers' professional preparation. The three review studies do, however, have some flaws. For instance, Coombe et al. (2020) did not specify the method used to choose the study subjects. Additionally, the study's concepts were applied somewhat inconsistently, with a concentration on evaluating literacy in general education as opposed to LAL. On the other hand, Lee & Butler (2020) described their review methods in depth. Other crucial components, such research designs or contexts, were left out as they reviewed the LAL studies' participants and research methodologies. The review study by Xu & Brown (2016) is thorough for teacher assessment literacy in general education, but not for teachers' LAL. In conclusion, additional study is required because the previous review studies have not offered a comprehensive picture of LAL.

The aim of this paper is to examine the research literature on assessment literacy in language education through review and analysis. For this purpose, this bibliometric review addressed the following research questions.

1. What are the volume, growth trajectory, and geographic distribution of scholarship on LAL researches?
2. Which scholars have emerged as thought leaders in the LAL literature?
3. What topics in LAL research have received the greatest attention in the literature?
4. What are the research gaps and potential future research directions for LAL?

The existing LAL literature primarily focuses on empirical and theoretical aspects, with limited systematic reviews. Such reviews are necessary due to the evolving and complex nature of LAL research. LAL differs from general assessment literacy (AL) and requires separate study. Using bibliometric can provide a more precise and efficient review of the expanding LAL literature, which lacks a consensus on an optimal theoretical framework due to its contextual and stakeholder-driven complexity. This study is new because, in our examination of the available literature data, we were unable to locate any literature review papers that specifically addressed AL in language instruction. As a result, this study offers perspectives that prior studies on this subject have not yet completely analyzed or documented. We searched the Scopus database for publications, which is one of the largest academic databases worldwide, that matched the research goals in order to answer these questions. To further enhance the understanding of the current research landscape, we conducted a systematic literature review and conducted a comprehensive analysis of the most recent state-of-the-art literature AL within the field of language education. This review encompassed publications from January 1994 to December 2021.

3. Methodology

3.1. Bibliometric Analysis

The proliferation of literature has made it challenging to organize and present studies in a clear and orderly manner. As a result, quickly and accurately identifying the essential literature related to a research topic has always been a perplexing task (Aria & Cuccurullo, 2017). Bibliometric analysis, which involves quantitatively analyzing articles published within a specific field (Baker et al., 2019), is one of the methods used for reviewing literature. It is a widely used approach for accumulating knowledge and structuring previous research findings, especially when examining various aspects of science and global rankings of institutions and universities (Ellegaard & Wallin, 2015). In response to this issue, co-word analysis, a type of bibliometric

technique, was developed to calculate the co-occurrence frequencies of chosen words in literature (Callon et al., 1991). In this study, we adopted a series of co-word network analyses to visually represent the network structure of pivotal keywords based on their co-occurrence relationships.

3.2. Data

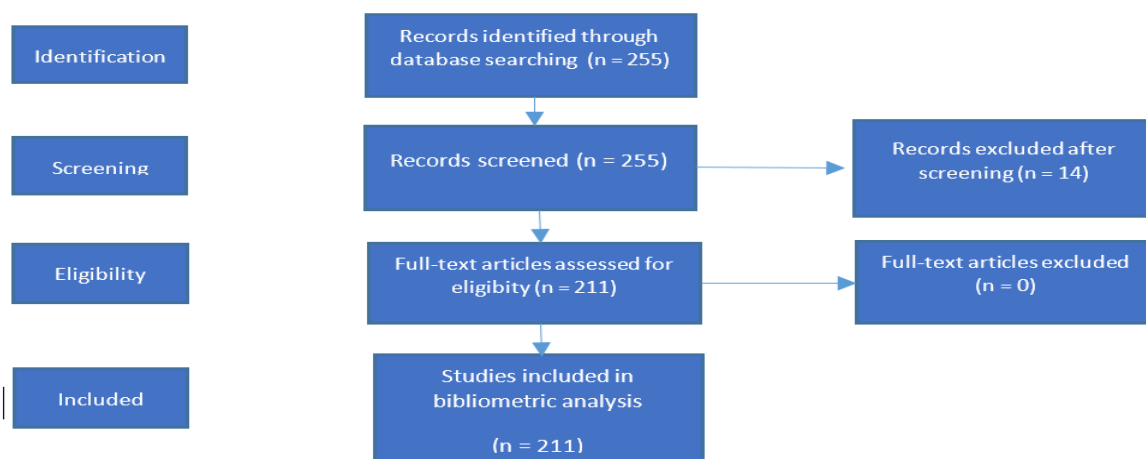
This research was based on Scopus database, which is one of the most significant scientific databases. To identify the articles that related to our topic, we focused on two keywords: “assessment literacy” and “language education” based on the aforementioned database and screening criteria. However, the specific keywords used in the search query can impact the resulting sample and subsequent analysis, and to yield new insights, alternative keywords were chosen for search queries. Specifically, apart from “assessment literacy”, the following relevant keywords were selected: “assessment knowledge”, “assessment practice*”, “assessment competence*”. Regarding language education, the alternatives “language teaching*”, “language learning*”, “language assessment” were included in the search queries. With regard to timeline, we terminated our search period at the end of 2021. We only selected publications written in English. Thus, the following search query was used to derive a primary database from Scopus:

TITLE-ABS-KEY (((“assessment literacy”) OR (“assessment knowledge”) OR (“assessment practice*”) OR (“assessment competence*”)) AND (“language teaching”) OR (“language learning”) OR (“language education”) OR (“language assessment”))) AND (EXCLUDE (PUBYEAR, 2023) OR EXCLUDE (PUBYEAR, 2022)) AND (LIMIT-TO (LANGUAGE, “English”)) AND (LIMIT-TO (SUBJAREA, “SOCI”))

The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) (Figure 1) was applied to ensure the quality of the document search processing (Moher et al., 2009). Our search yielded an initial 225 documents. We subsequently excluded 14 documents due to the full-text unavailability after screening based on the documents’ content. All duplicated records and conference papers were also removed. After this round of manual screen, 211 documents, including journal articles and books/book chapters, were obtained for final analysis. The following fields relating to each document were recorded in the final data: article identity number, article title, source journal, cited relationship, authors, institution, country, link of the document, and publication year.

Figure 1

PRISMA Diagram Identifying Procedure to Refine Documents for Bibliometric Analysis



3.3. Data Analysis

The “data” analyzed for this review consisted of bibliographic information describing features of the 211 Scopus-indexed documents. These “meta-data” included the author names, titles, publication dates, and author affiliations of the documents, as well as copious citation information. Descriptive statistics were used to conduct trend analyses related to the growth and geography of the LAL literature. A co-occurrence analysis of scientific domains cited in the literature was conducted to uncover the underlying structure of the research field’s scientific domains in the context of LAL. This analysis was facilitated using VOS viewer software, with a focus on making the scientific domain co-occurrence analysis more understandable.

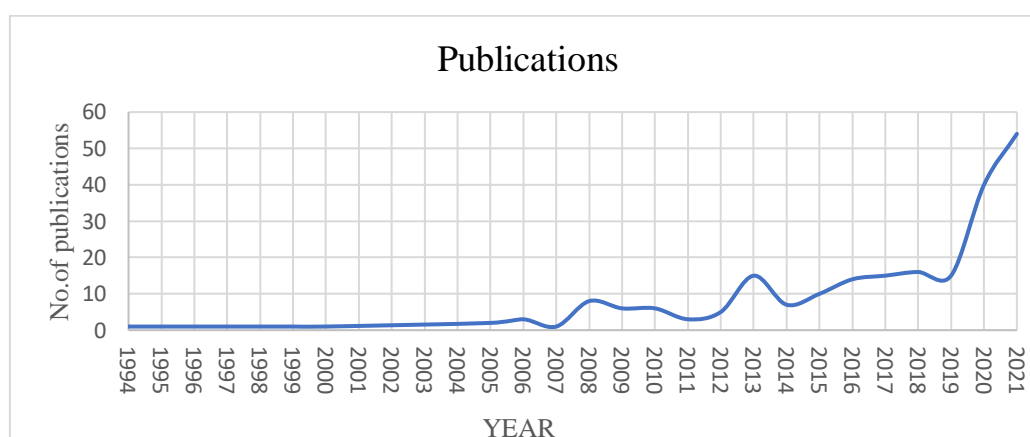
4. Results and Discussion

4.1. Publications and Citations

Figure 2 indicates the number of publications and citations developing between 1994 and 2021 from 211 papers.

Figure 2

Annual Growth of Publications on AL in Language Education Between 1994 and 2021

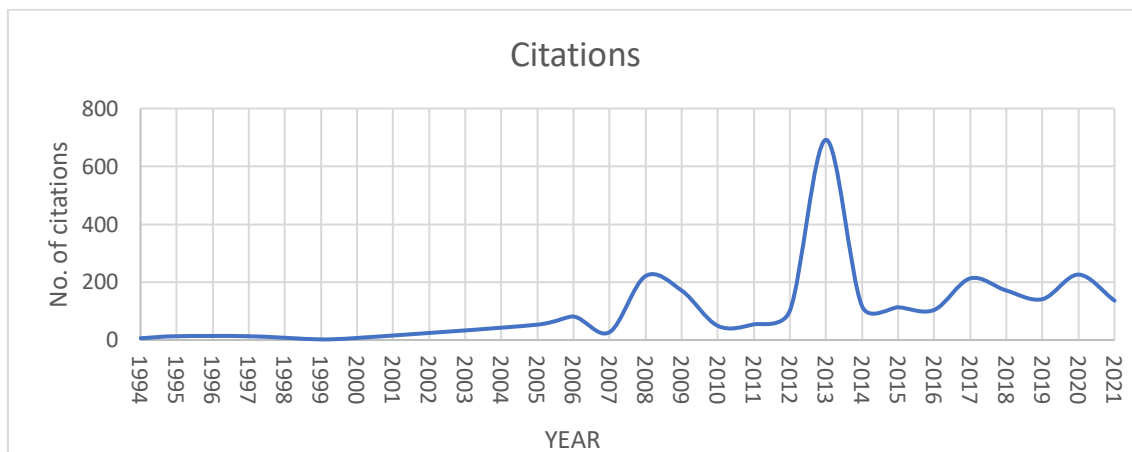


Based on the growth trends in publications on AL in language education between 1994 and 2021, we divided the publications into three periods:

- + 1994 – 2007: ignorant period, during which AL in language education seemed to be overlooked by scholars and only 11 documents were published (0.4% of the total publications);
- + 2008 – 2014: emergent period, during which AL in language education started to attract some attention from scholars and 50 documents were published (22.2% of the total publications);
- + 2015–2021: growing period, during which AL in language education received significant attention from scholars: 164 documents were published (72.9% of the total publications).

Figure 3

Annual Growth of Publications and Citations on AL in Language Education Between 1994 and 2021

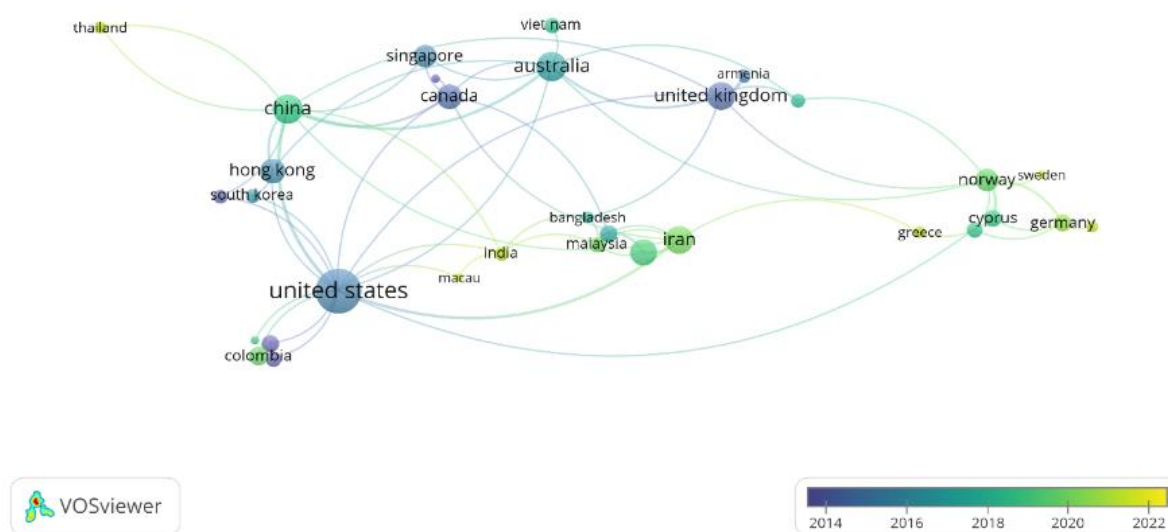


The first article in this field was published in 1994, increased horizontally to 2007 and fluctuated between 2007 and 2012. By 2013, the maximum number of publications reached a high of 692 citations. The number of citations decreased significantly after 2013 to 2016 with the number of citations being 104. However, from 2016 to 2021, the number of publications tended to fluctuate between 136 and 226.

Figure 4 shows the geographical contribution to AL in language education. Authors from 49 countries have (co)authored at least one publication on this topic. Surprisingly, three out-of-region countries, including the US (46 publications), Australia and China (17 publications each) were found to be the most productive countries beside three countries the UK and Iran (15 publications each) and Turkey (13 publications). These countries contributed nearly half of the total published documents. In addition, Canada, Norway and Singapore are countries that have also made a significant contribution to this topic.

Figure 4

The Geographical Contribution to AL in Language Education



4.2. The Most Prominent Authors, Publications and Outlets

Many AL-related documents in language education were co-authored by authors from the US countries, the UK, and Australia in collaboration with colleagues in other countries. Overall, there were 204 authors who participated in the production of papers on AL in language education.

Table 1

Top 15 Most Cited Authors by Total Articles and Citations

Rank	Authors	Affiliations	Cited by
1	Inbar-Lourie, O.	School of Education, Tel-Aviv University, Tel-Aviv 69978, Israel	132
2	Lee, I.	Faculty of Education, The Chinese University of Hong Kong, Hong Kong	118
3	Scarino, A.	Research Centre for Languages and Cultures, School of Communication, International Studies and Languages, University of South Australia, Magill Campus, Adelaide, Australia	105
4	Antón, M.	Indiana University, Purdue University, Indianapolis, IN, United States	84
5	Taylor, L.	Centre for Research in English Language Learning and Assessment (CRELLA), University of Bedfordshire, 47 Montague Road, Cambridge, CB4 1BU, United Kingdom	82
6	Malone, M.E.	Center for Applied Linguistics, Associate Vice President World Languages and International Programs, 4646 40th Street NW, Washington, DC 20016-1859, United States	72
7	Lam, R.	Hong Kong Baptist University, China	71
8	Leung, C. & Lewkowicz, J.O.	King's College, London, United Kingdom; American University of Armenia, Yerevan, Armenia	59
9	Pill, J. & Harding, L.	Language Testing Research Centre, The University of Melbourne, Babel Building, Parkville, VIC 3010, Australia; Lancaster University, United Kingdom	57
10	Hill, K. & McNamara T.	Medical Education Unit, School of Medicine, University of Melbourne, Victoria 3010, Australia	54
11	Bailey, A.L. & Heritage, M.	University of California, Los Angeles, CA, United States; National Center for Research on Evaluation, Standards, and Student Testing (CRESST), Los Angeles, CA, United States	48
12	McNamara, T.	School of Languages and Linguistics, University of Melbourne, VIC 3010, Australia	46
13	Gould, J.	University of South Australia, Adelaide, SA, Australia	43
14	Giraldo, F.	Universidad de Caldas, Manizales, Colombia	42

15	Barrot, J.S.	Department of English, National University, Manila, Philippines	41
----	--------------	---	----

Table 1 lists the top 15 most cited authors by total articles and citations according to the Scopus database. Most of the authors were from developed countries. Of the top 15 authors, only one (Inbar-Lourie, O. from Tel-Aviv University, Israel) received the largest number of citations, at 132, and 5 authors received more than 40 citations. The second most cited author was Lee, I. from Chinese University of Hong Kong, Hong Kong, with 118 citations.

Next, we tried to identify the most impactful articles published in this field. Table 2 presents 10 documents with the highest numbers of citations from 1994 to 2021. Notably, most of these articles were published after 2013, coinciding with the period of heightened citation activity. This suggests that these articles likely made significant contributions to the evolving landscape of this field, as they garnered substantial attention and recognition during a time of notable transformation and expansion. These ten articles discuss various topics such as Constructing a language assessment knowledge base (Inbar-Lourie O., 2008), Classroom writing assessment and feedback (Lee, I., 2017) and Language assessment literacy as self-awareness (Scarino, A., 2013). Most articles present qualitative studies and some use quantitative methodologies. The 10 most common articles which are commonly referred to by other researchers are listed in Table 2 following.

Table 2

Top Ten Most Cited Articles

Authors	Title	Year	Source title	Cited by	Design
Inbar-Lourie, O.	Constructing a language assessment knowledge base: A focus on language assessment courses	2008	Language Testing	132	Qualitative
Lee, I.	Classroom writing assessment and feedback in L2 school contexts	2017	Classroom Writing Assessment and Feedback in L2 School Contexts	118	Qualitative (Book)
Scarino, A.	Language assessment literacy as self-awareness: Understanding the role of interpretation in assessment and in teacher learning	2013	Language Testing	105	Qualitative
Antón, M.	Dynamic assessment of advanced second language learners	2009	Foreign Language Annals	84	Qualitative
Taylor, L.	Communicating the theory, practice and principles of language testing to test stakeholders: Some reflections	2013	Language Testing	82	Qualitative
Malone, M.E.	The essentials of assessment literacy: Contrasts between testers and users	2013	Language Testing	72	Qualitative
Lam, R.	Language assessment training in Hong Kong: Implications for	2015	Language Testing	71	Qualitative

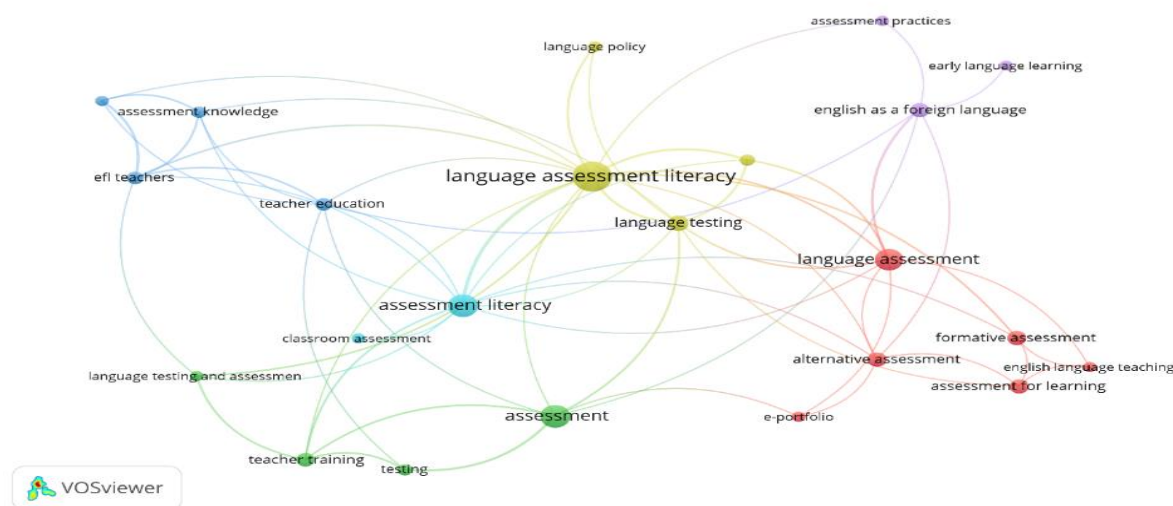
language assessment literacy					
Leung, C., Lewkowicz, J.O.	Expanding horizons and unresolved conundrums: Language testing and assessment	2006	TESOL Quarterly	59	Qualitative
Pill, J., Harding, L.	Defining the language assessment literacy gap: Evidence from a parliamentary inquiry	2013	Language Testing	57	Qualitative
Hill K., McNamar a, T.	Developing a comprehensive, empirically based research framework for classroom-based assessment	2012	Language Testing	54	Qualitative

4.2. Documents Analysis Co-occurrence by Keywords

The objective of the first stage of the analysis was to explore topics related to AL in language education in research. A co-occurrence analysis of keywords was conducted using VOSviewer on a dataset of 211 documents, aiming to identify common themes and track the evolution of research topics in the field of LAL. Keeping 3 as the minimum number of author keywords occurrences within the identified articles, 24 out of 452 keywords met the threshold. For each of the remaining 24 keywords, the total strength of the co-occurrence links with other keywords was calculated by the VOSviewer software using normalization of associations and full counting algorithm. VOSviewer output is a network (Fig. 5) composed of 24 nodes corresponding to 6 clusters.

Figure 5

Co-occurrence by Keywords of AL in Language Education



The analysis revealed that LAL studies primarily revolve around key themes such as “assessment literacy,” “language assessment literacy,” “language assessment,” “teacher education,” “teacher training,” and “EFL.” These themes indicate that LAL research is concentrated around topics related to language assessment, particularly concerning teachers in school and classroom settings. Furthermore, certain themes, such as “assessment literacy,” “language assessment literacy,” “language assessment,” held central positions in the network,

indicating their importance and prevalence in LAL research. “Teacher education,” “teacher training,” and “assessment practices” also experienced similar growth during this period. This suggests that LAL research has diversified its focus while maintaining core themes, with a significant emphasis on the role of teachers as key stakeholders implementing LAL practices.

Based on the VOS viewer output, clusters were named based on the topics researched, such as “language assessment literacy,” “language assessment,” “teacher education,” “teacher training,” “EFL, ” and “classroom assessment.” In the following sub-sections, the obtained clusters are described through some relevant works on the subject to outline the research interests related to AL in language education.

+ Cluster 1: Language assessment literacy

According to Figure 5, the highest number of occurrences based on keywords are: “language assessment literacy”, “language testing” and “language policy” which show a very strong relationship between these three keywords in the topic of research. This is further proven by the total link strength of 8, and 13 link strength for “language assessment literacy” and “language policy”. Also, this largest cluster appears bright yellow, indicating that the studies in the cluster are newer. It indicates that LAL has been recently researched in accordance with educational policies of language teaching.

+ Cluster 2: Language assessment

In the second cluster, “alternative assessment” and “formative assessment” had a higher frequency. The timeline of keyword co-occurrence nodes reveals the evolution and refinement of LAL topics over time. In terms of the number of occurrences, size of alternative assessment is quite big (7) compared to other terms in this clusters (Figure 5). To be more detailed, specific assessment methods like “formative assessment”, “e-portfolio” occurred around 3 in this cluster, showcasing the development of detailed assessment approaches to enhance language assessment knowledge and practices.

+ Cluster 3: Teacher education

In the third cluster, the term “EFL teachers” was the most frequently mentioned, appearing 11 times. It occupies a central position and exhibits the highest strength with 11 links. Within this cluster, the term “teacher education” was found to have a close association with “assessment knowledge,” with respective link strengths of 9 and 7. This suggests that there is increased emphasis on the assessment literacy of EFL teachers in comparison to general assessment practices, and there is a growing focus on whether teacher education programs are effective in cultivating assessment knowledge.

+ Cluster 4: Teacher training

In the fourth cluster, two terms namely “teacher training” and “language assessment and testing” were found to be closely related compared with other terms. These two terms also have a big circle indicating that they are dominant terms in this cluster, with each represented by 6 and 4 total number of occurrences respectively. Some of the terms such as testing and assessment are scattered and not closely linked. This observation implies that numerous researchers have undertaken investigations concerning the state of LAL among educators and have delved into their training needs. Consequently, it underscores the considerable scholarly focus on LAL training for teachers within the academic community.

+ Cluster 5: EFL

In the fifth cluster, the terms “English as a foreign language” had the biggest circle with a total of 8 occurrences and 7 total links of strength. The term “early language learning” is the

second highest in this cluster, demonstrating researchers' significant interest in foreign language teachers when exploring LAL. Also, there are some terms that "assessment practices" have been done since the early stage in the language education.

+ Cluster 6: Classroom assessment

In the sixth cluster, 2 occurrences were seen in classroom assessment with assessment literacy. The circle of the term classroom assessment is smaller than assessment literacy represented by 19 and 17 total number of occurrences and total link of strength respectively. These findings reflect that the issue of classroom assessment is of much concern by researchers around the world.

5. Discussion on Research Gaps and Future Research Directions

The study's conclusions showed that from 1994 till 2021, improvements in research and publishing have been made in the area of assessment literacy in language education. This is in line with the claim that LAL research has developed into a thriving research subfield in language testing and assessment and is no longer "*in its infancy*" as Fulcher (2012, p. 117) claimed more than a decade ago (Harding & Brunfaut, 2020). Numerous LAL studies have been published in high-impact language testing and assessment publications, which serves as one example of the emphasis on the field. Significant articles have also resulted from conferences or symposiums hosted by organizations with LAL as the focus as well as LAL special issues of these journals (such as Language Testing 2013). The results also demonstrated how several research in the social sciences, notably in the teaching and learning of languages, have examined literacy assessments. This can be as a result of the fact that this area of expertise focuses on understanding educators, teachers, and students.

We also discovered that LAL studies were under-explored in the contexts of Africa and Latin America. This data lends some credence to the claims made about the geographic origins of LAL study participants by Harding and Kremmel (2016) and Tsagari (2020). This may suggest that LAL hasn't received as much attention as it should in these regions. Instead, LAL research was most common in the Asia-Pacific region, Europe, and the Middle East. This outcome is not unexpected given that evaluation activities are typically performed as a result of the high value placed on EFL training in these situations.

The findings also demonstrated that in the domain of language teaching, qualitative research designs dominated AL research. Instead of using quantitative designs, the majority of studies used mixed-methods and qualitative ones. It can conclude that there is not enough psychometric evidence to warrant the assessment of literacy. While the majority of studies focused on language teachers, particularly EFL teachers, very few studies were conducted from the perspectives of students, policymakers, language testers, teacher educators, and other stakeholders.

This study revealed a proliferation of research that focuses on language instructors' LAL, particularly EFL teachers, and a lack of concern for other stakeholders. These results support the assertions made in various studies (e.g., Lee & Butler, 2020; Pill & Harding, 2013). Language teachers acquire LAL most and need to advance their own LAL before imparting it to learners because they are the main stakeholder who performs a variety of assessment activities inside and outside of the classrooms (Vogt, Tsagari, & Csépes, 2020). Language teachers have so continually received attention in LAL scholarship, which is not surprising.

Other stakeholders, such as policymakers, language testers, admissions officers, test

developers, teacher educators, administrators, etc., are underrepresented in LAL studies in addition to learners. Other stakeholders among them are deemed to be removed from the primary assessment operations, with the exception of language testers and test developers (Taylor, 2013). Their LAL is supposedly less noticeable than that of teachers, students, language testers, or test creators who are directly subjected to testing and assessment. Actually, these stakeholder groups must also obtain the proper LAL in accordance with the extent of their participation in assessment activities occurring outside of classrooms (Yan & Fan, 2021). As a result, greater scholarly focus should be given to what LAL is required and how LAL is generated among stakeholders at the assessment core and peripheral.

6. Conclusion

The current study has identified four key findings from a review of LAL studies from peer-reviewed journal articles, conference papers, and book chapters:

(1) LAL studies have been on the rise and become a thriving subfield in language testing and assessment;

(2) most LAL studies were contextualised in the Asia-Pacific region, European countries and the Middle East;

(3) existing studies preferred qualitative designs over quantitative and mixed-methods research;

(4) an overwhelming majority of studies focused on language teachers, especially EFL teachers, while few studies were conducted from perspectives of learners, policy makers, language testers and other stakeholders;

The analysis revealed that LAL's theoretical framework is currently under construction, with a notable emphasis on the perspectives of key stakeholders, particularly “teachers” Research on other stakeholder groups, especially students, has also gained prominence. Future research should explore diverse stakeholders in LAL.

Our review has offered suggestions for future LAL research based on the findings. To acquire a more comprehensive understanding of LAL, it is first important to incorporate the opinions of more professional stakeholders. Expanded research is warranted on LAL of language teachers other than EFL teachers at various career stages. Second, the scope of LAL research should be widened to include subtler methods. More longitudinal research is desired to investigate the LAL developmental trajectories of stakeholders using information gathered through think aloud, journal writing, and narrative framing. Third, LAL regionalization is a developing area of study that requires adequate consideration. It is advised to list and examine common characteristics of LAL development in various circumstances.

The aforementioned findings are anticipated to alert researchers to fresh directions in LAL study and refocus their focus on unexplored areas in LAL literature. The difficulties that academics in this field of knowledge frequently raise need to be given more attention by stakeholders and employers. To meet the needs of more targeted training, a proposed systematic assessment can be carried out to determine the level of language assessment literacy and the elements that influence it. Finally, LAL has emerged as a promising study direction in language testing and evaluation. When contemplating the aforementioned potential research directions, academics are urged to look outside the box.

The research method employed in this study has certain limitations. Firstly, the data collection was limited to Scopus while the other databases include influential journals, valuable

publications from other databases may have been omitted. Furthermore, the study focused exclusively on English language journal articles, excluding non-English articles, dissertations, and conference papers. Additionally, to ensure the precision of search results, particularly in collecting and analyzing articles highly centered on LAL, a more stringent search approach was adopted, potentially resulting in a lower number of search results. Future research should consider expanding the scope of the database search and incorporating multiple languages to provide a more comprehensive understanding of developments in the field of LAL, thereby enhancing the richness of the visualized data.

References

- Aria M., Cuccurullo C. (2017). Bibliometrix : An R-tool for comprehensive science mapping analysis. *Journal of informetrics*, 11(4), 959-975. <https://doi.org/10.1016/j.joi.2017.08.007>
- Baker, B. A., & Riches, C. (2018). The development of EFL examinations in Haiti: Collaboration and language assessment literacy development. *Language Testing*, 35(4), 557–581. <https://doi.org/10.1177/0265532217716732>
- Baker HK., Kumar S, Pandey N. (2020). A bibliometric analysis of managerial finance: a retrospective. *Managerial Finance*. <https://doi.org/10.1108/MF-06-2019-0277>
- Kremmel, B., & Harding, L. (2020). Towards a Comprehensive, Empirical Model of Language Assessment Literacy across Stakeholder Groups: Developing the Language Assessment Literacy Survey. *Language Assessment Quarterly*, 17(1), 100-120. <https://doi.org/10.1080/15434303.2019.1674855>
- Black, P., & Wiliam, D. (2018). Classroom assessment and pedagogy. *Assessment in Education: Principles, Policy & Practice*, 25(6), 551–575. <https://doi.org/10.1080/0969594X.2018.1441807>
- Brindley, G. (2001). Language assessment and professional development. In Elder, C., Brown, A., Grove, E., Hall, K., Iwashita, N., Lumley, T., McNamara, T., and O’Loughlin, K. (Eds.) *Experimenting with Uncertainty. Essays in Honour of Alan Davies*. Cambridge: Cambridge University Press, 126–136.
- Brown, H. D., & Abeywickrama, P. (2010). *Language assessment: Principles and classroom practices*. New York: Pearson Longman.
- Callon, M., Law, J., & Rip, A. (1986). *Mapping the dynamics of science and technology: Sociology of science in the real world* (M. Callon, J. Law, & A. Rip, Eds.) Palgrave Macmillan UK. <https://doi.org/10.1007/978-1-349-07408-2>
- Cheng, L., Rogers, T., & Hu, H. (2004). ESL/EFL instructors’ classroom assessment practices: Purposes, methods and procedures. *Language Testing*, 21(3), 360–389. <https://doi.org/10.1191/0265532204lt288oa>
- Coombe, C., Vafadar, H., & Mohebbi, H. (2020). Language assessment literacy: What do we need to learn, unlearn, and relearn? *Language Testing in Asia*, 10. Article 3. <https://doi.org/10.1186/s40468-020-00101-6>
- Davies, A. (2008). Textbook trends in teaching language testing. *Language Testing*, 25(3), 327–347. <https://doi.org/10.1177/0265532208090156>
- Deygers, B., & Malone, M. E. (2019). Language assessment literacy in university admission policies, or the dialogue that isn’t. *Language Testing*, 36(3), 347–368. <https://doi.org/10.1177/0265532219826390>
- Ellegaard O, Wallin, JA (2015). The bibliometric analysis of scholarly production: How great is the impact? *Scientometrics*, 105(3), 1809-1831. [10.1007/s11192-015-1645-z](https://doi.org/10.1007/s11192-015-1645-z)
- Fulcher, G. (2012). Assessment literacy for the language classroom. *Language Assessment Quarterly*, 9(2), 113–132. <https://doi.org/10.1080/15434303.2011.642041>
- Harding, L., & Brunfaut, T. (2020). Trajectories of language assessment literacy in a teacherresearcher partnership: Locating elements of praxis through narrative inquiry. In M. E. Poehner & O. Inbar-Lourie (Eds.), *Towards a re-conceptualization of second language classroom assessment* (pp. 61–81). Springer International Publishing. <https://doi.org/10.1007/978-3-030-35081-9>
- Harding, L., & Kremmel, B. (2016). Teacher assessment literacy and professional development. In D. Tsagari & J. Banerjee (Eds.), *Handbook of second language assessment* (pp. 413–427). De Gruyter Mouton. <https://doi.org/10.1515/9781614513827>

- Inbar-Lourie, O. (2008). Constructing a language assessment knowledge base: A focus on language assessment courses. *Language Testing*, 25(3), 385–402. <https://doi.org/10.1177/0265532208090158>
- Inbar-Lourie, O. (2013). Guest Editorial to the special issue on language assessment literacy. *Language Testing*, 30(3), 301–307. <https://doi.org/10.1177/0265532213480126>
- Inbar-Lourie, O. (2017). Language assessment literacy. In E. Shohamy, I. Or, & S. May (Eds.), *Language testing and assessment, encyclopaedia of language and education* (pp. 257–286). Springer International Publishing. https://doi.org/10.1007/978-3-319-02261-1_19
- Lee, J. Y., & Butler, Y. (2020). Reconceptualizing language assessment literacy: Where are language learners? *TESOL Quarterly*, 54(4), 1098–1111. <https://doi.org/10.1002/tesq.576>
- Malone, M. (2013). The essentials of assessment literacy: Contrasts between testers and users. *Language Testing*, 30(3), 329–344. https://doi.org/10.1007/978-3-319-02261-1_19
- Pill, J., & Harding, L. (2013). Defining the language assessment literacy gap: Evidence from a parliamentary inquiry. *Language Testing*, 30(3), 381–402. <https://doi.org/10.1177/0265532213480337>
- Taylor, L. (2013). Communicating the theory, practice and principles of language testing to test stakeholders: Some reflections. *Language Testing*, 30(3), 403–412. <https://doi.org/10.1177/0265532213480338>
- van Eck, N.J., Waltman, L. (2014). Visualizing Bibliometric Networks. In: Ding, Y., Rousseau, R., Wolfram, D. (eds), *Measuring Scholarly Impact*. Springer, Cham. https://doi.org/10.1007/978-3-319-10377-8_13
- Vogt, K., Tsagari, D., & Csépes, I. (2020). Linking learners' perspectives on language assessment practices to Teachers' Assessment Literacy Enhancement (TALE): Insights from four European countries. *Language Assessment Quarterly*, 17(4), 410–433. <https://doi.org/10.1080/15434303.2020.1776714>
- Xu, Y., & Brown, G. T. L. (2016). Teacher assessment literacy in practice: A reconceptualization. *Teaching and Teacher Education*, 58, 149–162. <https://doi.org/10.1016/j.tate.2016.05.010>
- Yan, X., Zhang, C., & Fan, J. S. (2018). “Assessment knowledge is important, but . . . ”: How contextual and experiential factors mediate assessment practice and training needs of language teachers. *System*, 74, 158–168. <https://doi.org/10.1016/j.system.2018.03.003>

NĂNG LỰC KHẢO THÍ TRONG GIẢNG DẠY NGÔN NGỮ: CÂU CHUYỆN ĐÃ KỂ CHO ĐẾN NAY

Nguyễn Thị Tố Loan¹, Dương Thu Mai²

¹Khoa Ngoại ngữ, Trường Đại học Hùng Vương, Nông Trang, Việt Trì, Phú Thọ, Việt Nam

²Khoa Sư phạm tiếng Anh, Trường Đại học Ngoại ngữ, Đại học Quốc gia Hà Nội,
Đường Phạm Văn Đồng, Cầu Giấy, Hà Nội, Việt Nam

Tóm tắt: Trong ba thập kỷ qua, đã có sự gia tăng đáng kể trong nghiên cứu về năng lực khảo thí trong giảng dạy ngôn ngữ hay còn gọi là năng lực khảo thí ngôn ngữ (viết tắt là LAL), thể hiện bằng sự quan tâm ngày càng lớn của các nhà nghiên cứu hàn lâm và các xuất bản phẩm mới. Do tầm quan trọng của LAL đối với sự phát triển chuyên môn của giáo viên ngôn ngữ, bài viết này áp dụng một quy trình đánh giá có hệ thống để trình bày một đánh giá hiện trạng tổng quan về LAL. Trên dữ liệu của 211 xuất bản phẩm có liên quan được xác định bằng công cụ tìm kiếm Scopus, dữ liệu được phân tích bằng phần mềm VOSviewer để tạo ra phân tích trắc lượng thư mục. Nghiên cứu này đã đưa ra những quan điểm chưa được khám phá hoặc đánh giá đúng đắn bởi các nghiên cứu khác về chủ đề này. Các phát hiện này có một số ý nghĩa liên quan đến tổng quan về LAL hiện nay, chỉ ra các hướng nghiên cứu trong tương lai cũng như các lỗ hổng nghiên cứu. Kết quả của nghiên cứu cung cấp một khuôn khổ vững chắc để hiểu thấu đáo hơn về sự phát triển của các chủ đề nghiên cứu, phương pháp nghiên cứu và xu hướng trong lĩnh vực nghiên cứu khá mới mẻ và hấp dẫn này.

Từ khóa: giáo dục ngôn ngữ, năng lực khảo thí, phân tích trắc lượng thư mục, nghiên cứu hệ thống