

# An Analysis of Research Progress and Frontiers of Integrated Language Assessment Using Citespace

Lan Luo

Guangxi University of Foreign Languages, Nanning, Guangxi, China  
nie19.ll@e.ntu.edu.sg

## ABSTRACT

The purpose of the current research is to explore the research development trend and hotspots of integrated assessment in the field of language. Using CiteSpace, a Java application for visualizing trends in scientific literature, this paper analyzed 512 publications between 2000 and 2020 from Web of Science, Scopus, and Google Scholar. By setting node types as “keyword” and “reference”, respectively, we have identified the development trends and the interests of researchers in the related field based on the visual knowledge map generated by Citespace. The study indicates that researchers have transferred from the traditional concept of “summary writing” to an overarching phrase “integrated writing” and “integrated task”. They also tend to incorporate multi-modalities in the past decade. Besides, researchers in the field mainly focus on computer-assisted language learning, assessment task, cross-disciplinary academic language skills, and theories of integrated language assessment. The work makes both theoretical and methodological contributions to progressive integrated language theory building, assessment task design, and language learning.

## Keywords:

Citespace; integrated assessment; language testing

Article Received: 18 October 2020, Revised: 3 November 2020, Accepted: 24 December 2020

## INTRODUCTION

Studies in integrated assessment for second language (L2) learners have received much attention in academic settings in the past decades. Many language tests (e.g., PTE Academic Test, Test of English as a Foreign Language (TOEFL), and Trinity's Integrated Skills in English) have included integrated tasks in their English for academic purpose assessment batteries. The rationale behind these initiatives is that the integrated task simulates the target language use domain, its score can be extrapolated to authentic academic contexts [1,2,3], and it attenuates the negative effects of test-takers' differences in their background knowledge [3, 4].

Nevertheless, this topic has posed new challenges to language assessment because different researchers have different focus, depending on their research interests. In contexts like academic settings, many researchers [2, 5] have defined integrated assessment as the integration of macro-skills (e.g., listening-to-writing, or reading-listening-writing), which are

essential channels in authentic English-medium institutions of education. Specifically, these researchers further narrowed their focus to a strong version of skills integration (i.e., giving opinion by citing source materials to support one's position) [6] or a weaker or middle-range version (i.e., answering questions by using sources or summarizing source materials) [7, 8]. Other scholars have paid attention to the integration of language and content subject [9, 10], the integration of the use of technology for task completion or automated scoring [11], or the use of different sources of input (i.e., text, audio, video) [12] as part of the construct of language assessment. Though such variations in definition and operationalization are healthy for developing the field itself [13], further research would be restricted because of this lack of robust, coherent definition as well as a lack of a systematic review of the development progress of the integrated assessment. Recently, scientometrics relying on databases has been increasingly used in social science research. It enables researchers to extract

necessary information to work on their areas of interest in a short time. To inspire future research, the current study used Citespace 5.7.R2 to analyze the research trends in the past two decades and hotspots in the field.

## MATERIALS AND METHODS

### A. Data Sources and Retrieval Strategy

A pilot study showed that integrated assessment has been increasingly investigated since 2000; therefore, we set the timespan from 2000 to 2020. Related publication regarding integrated assessment was retrieved through the Web of Science Core Collection, Scopus, and Google Scholar on November 15, 2020. The search expressions were constructed as follows: TITLE-ABS-KEY= ( "integrated language assessment" ) OR ( "integrated assessment" AND "language" ) OR ( "integrated language task" ) OR ( "integrated task" AND "language" ) OR ( "integrated skill" AND "language" ) OR ( "skills integration" AND "language" ) OR ( "integrated language test" ) OR ( "integrated test" AND "language" ) OR ( "integrative test" AND language ) OR ( "integrative task" AND language ) OR ( "integrated reading" ) OR ( "integrated writing" ) OR ( "integrated speaking" ) OR ( "integrated listening" ) OR ( "writing from sources" ) OR ( "source-based writing" ) OR ( "content-responsive writing" ) OR ( "listening to writing" ) OR ( "listening into writing" ) OR ( "reading to writing" ) OR ( "reading into writing" ) OR ( "writing to reading" ) OR ( "writing into reading" ) OR ( "listening to speaking" ) OR ( "listening into speaking" ) OR ( "listening-reading-writing" ) OR ( "reading-listening-writing" ) OR ( "listening-reading-speaking" ) OR ( "reading-listening-speaking" ) OR ( "oral\* summar\*" ) OR ( "summary task" AND language ) OR ( "summary writing" ) OR ( "Discourse

synthesis" AND "language" ) OR ( "speaking summar\*" ) OR ( "writing summar\*" ) ) AND Language=English AND Time range=2000-2020. The initial search yielded 1873 publications between 2000 and 2020. The author and an assistant retrieved, downloaded, extracted, and removed duplicate data independently, and then discussed and standardized any controversies. After data cleansing, there remained 512 articles, which were imported into Citespace 5.7.R2 for data processing.

### B. Parameter Settings and Data Analysis

We have used several parameters that were recommended by the developer of Citespace [14]: a duration of 1 year was set for every time slice; the threshold level was Top 50; and the clipping mode was Pathfinder. To explore the evolvement and new trends, keyword was set as the node type; to investigate the research frontiers, reference was set as the node type because highly cited references could be clustered in terms of their similarities and interconnectivities [15]. Each cluster represented a specific knowledge domain. This documentation cluster analysis could be used to detect the research frontiers.

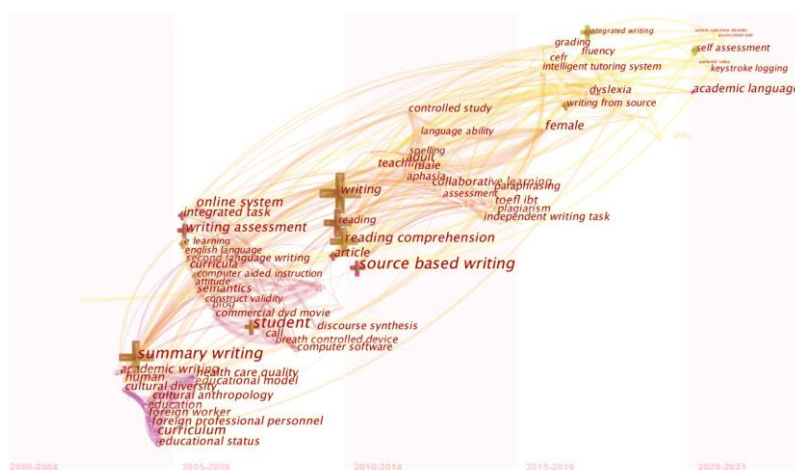
## RESULTS

### C. Development Trend of Integrated Assessment

As shown in Figure 1, a keyword time-zone view was generated. By synthesizing the time-zone map as well as the keyword frequency, centrality, year, burst, and time-zone map generated by Citespace, we sliced the development of integrated assessment into four time periods: (1) Before 2004, much of the research seemed focused on “summary writing” or “academic writing in the classroom”. However, since 2004, the appearance of the term “integrated task” indicated a paradigm shift from purely the integration of reading-into-writing to the

integration of more language skills, including listening and speaking. Much critical attention was also directed to computer-assisted language instruction and the integration of technology and language testing. For example, [16] has used key-logging to examine test-takers' revision behavior when they respond to an integrated writing task on a computer. [17] have also developed three computer-administered integrated assessment tasks to understand learners' task motivation. (2) The year 2010 marked a watershed in language assessment history in that the academia embarked on the research of "source-based writing", which meant that the past integrated writing began to incorporate expressing opinions in addition to summarizing ability as a construct of integrated writing tasks; (3) Since 2014, there has been a substantial increase in interest in comparing

integrated tasks with independent tasks by using high-stakes assessment tools like the TOEFL iBT. There has been an ongoing debate in the literature as to whether scores in independent tasks are good predictors to those in integrated tasks; (4) From 2018 to 2020, several keywords as “dyslexia”, “self-assessment”, and “keystroke logging” were highlighted. The integrated assessment was viewed as a tool offering diagnostic value in the field of neurocognition, education, and psychology. Keystroke logging and eye-tracking technology were used to investigate the writing processes of test-takers in order to build cognitive validity of language tests. Another highlighted term “CEFR”<sup>1</sup> indicated that the foci of the literature in this period was to map different integrated tasks onto a generally accepted framework.



**Figure 1: Keyword time-zone map of integrated assessment 2000 – 2020**

#### D. Research Frontier of Integrated Assessment

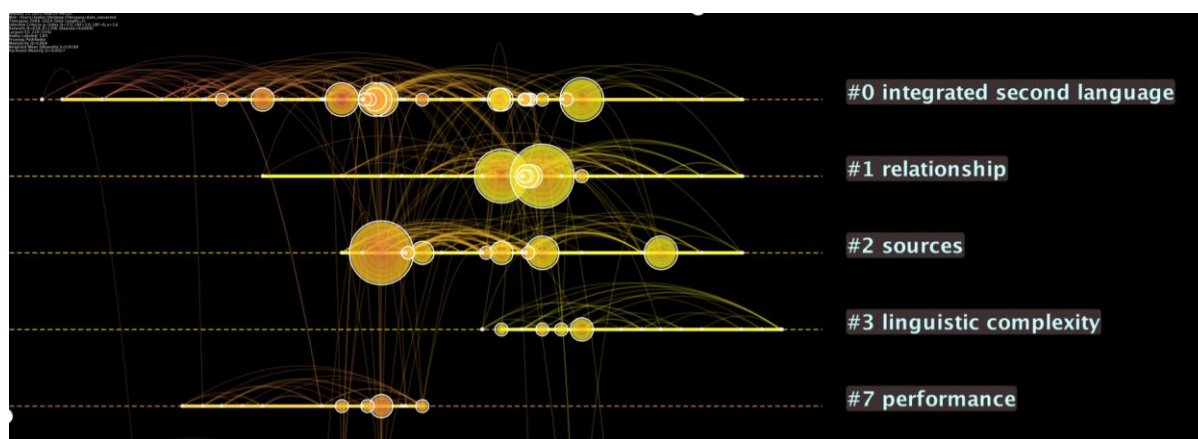
Figure 2 shows a timeline view of the co-citation network. #0 represents the largest cluster which contains the most publications. The bigger and the more nodes in a cluster, the larger the amount of attention it receives, thereby representing the research frontiers [14]. We have found that cluster #0, #1, and #2 contain the most nodes, indicating that they are the new trends of research.

As expected, cluster #0 represented the integrated nature of the topic in question; most publications

focused the integration of four macro language skills. The label “relationship” of cluster #1 can be interpreted from the three perspectives: (1) the connections between independent and integrated tasks. As a case in point, [18] have compared lexical sophistication in both tasks. (2) the relationship between sub-processes in discourse synthesis. For example, [19] have explored the relationship of organization and connection with scores in a reading-listening-writing task. (3) the relationship between different language skills and their contributions to the writing scores [20, 21].

Finally, cluster #2 indicated that source use in integrated tasks were brought to the fore. Researchers [22, 23, 24] have given a close investigation into writers' source use pattern; plagiarism; citation instruction; the functions the sources serve; and how discourse features, proficiency level, and task types influence discourse synthesis. It should be noted that some scholars [12] began to incorporate videos as sources of integrated writing. However, there was

an emerging trend that some researchers have shifted away from the traditional linear reading-to-writing towards a reciprocal integration of skills. For instance, [25, 26] have developed an iterative integrated task (writing-reading-writing) to provide insight into the complex processes linking reading and writing in L2. She argued that "Writing-into-reading can be a useful means to elicit skill integration processes."



**Figure 2. A timeline visualization of clusters of integrated assessment**

## CONCLUSIONS AND PROSPECTS

The current study uses Citespace to analyze the 2000-2020 English literature on integrated language assessment. It is found that the current research interest resides in integrated writing, which includes multiple modalities such as reading, listening, and writing as well as technology and videos; there is a lack of research on speaking as an output in integrated assessment, or the integration of subject content and language instruction, the void of which is expected to be filled by future research. In addition, the cross-disciplinary characteristic of integrated assessment is emphasized as researchers began to investigate it from neurological, psychological, and technological perspectives. By systematically summarizing the development process of integrated assessment research, this study lays a foundation for conducting relevant future research.

## REFERENCES

- [1] Chapelle, C. A., Enright, M. K., & Jamieson, J. M. (2008). Test score interpretation and use. Building a validity argument for the Test of English as a Foreign Language, 1-25.
- [2] Cumming, A. (2013). Assessing integrated skills. The companion to language assessment, 1, 216-229.
- [3] Read, J. (1990). Providing relevant content in an EAP writing test. English for specific purposes, 9(2), 109-121.
- [4] Weigle, S. C. (2004). Integrating reading and writing in a competency test for non-native speakers of English. Assessing writing, 9(1), 27-55.
- [5] Bozorgian, H. (2012). The relationship between listening and other language skills in international English language testing system. Theory and Practice in Language Studies, 2(4), 657-663.

- [6] Kibler, A. K., & Hardigree, C. (2017). Using evidence in L2 argumentative writing: A longitudinal case study across high school and university. *Language Learning*, 67(1), 75-109.
- [7] Keck, C. (2006). The use of paraphrase in summary writing: A comparison of L1 and L2 writers. *Journal of Second Language Writing*, 15(4), 261-278.
- [8] Yang, H.-C. (2014). Toward a model of strategies and summary writing performance. *Language Assessment Quarterly*, 11(4), 403-431.
- [9] Banegas, D. L. (2012). Integrating content and language in English language teaching in secondary education: Models, benefits, and challenges. *Studies in Second Language Learning and Teaching*, 2(1), 111-136.
- [10] Tedick, D. J., & Cammarata, L. (2012). Content and language integration in K-12 contexts: Student outcomes, teacher practices, and stakeholder perspectives. *Foreign Language Annals*, 45(s1), s28-s53.
- [11] Ramineni, C., Trapani, C. S., Williamson, D. M., Davey, T., & Bridgeman, B. (2012). Evaluation of the e-rater® Scoring Engine for the TOEFL® Independent and Integrated Prompts. ETS Research Report Series, 2012(1), i-51.
- [12] Machili, I., Papadopoulou, I., & Kantaridou, Z. (2020). Effect of strategy instruction on EFL students' video-mediated integrated writing performance. *Journal of Second Language Writing*, 48, 100708.
- [13] Yu, G. (2013). From integrative to integrated language assessment: Are we there yet? In: Taylor & Francis.
- [14] Chen, C. (2006). CiteSpace II: Detecting and visualizing emerging trends and transient patterns in scientific literature. *Journal of the American Society for information Science and Technology*, 57(3), 359-377.
- [15] Shi, Y., & Liu, X. (2019). Research on the literature of green building based on the Web of Science: a scientometric analysis in CiteSpace (2002-2018). *Sustainability*, 11(13), 3716.
- [16] Barkaoui, K. (2016). What and when second-language learners revise when responding to timed writing tasks on the computer: The roles of task type, second language proficiency, and keyboarding skills. *The Modern Language Journal*, 100(1), 320-340.
- [17] Kormos, J., Brunfaut, T., & Michel, M. (2020). Motivational Factors in Computer-administered Integrated Skills Tasks: A Study of Young Learners. *Language Assessment Quarterly*, 17(1), 43-59.
- [18] Kyle, K., & Crossley, S. (2016). The relationship between lexical sophistication and independent and source-based writing. *Journal of Second Language Writing*, 34, 12-24.
- [19] Plakans, L., & Gebril, A. (2017). Exploring the relationship of organization and connection with scores in integrated writing assessment. *Assessing writing*, 31, 98-112.
- [20] Cheong, C. M., Zhu, X., & Liao, X. (2018). Differences between the relationship of L1 learners' performance in integrated writing with both independent listening and independent reading cognitive skills. *Reading and Writing*, 31(4), 779-811.
- [21] Spivey, N. N. (2020). Discourse synthesis: creating texts from texts. In *Reading Empirical Research Studies* (pp. 469-509): Routledge.
- [22] Homayounzadeh, M., Saadat, M., & Ahmadi, A. (2019). Investigating the effect of source characteristics on task comparability in integrated writing tasks. *Assessing writing*, 41, 25-46.
- [23] Vandermeulen, N., Van den Broek, B., Van Steendam, E., & Rijlaarsdam, G. (2020). In search of an effective source use pattern for writing argumentative and informative synthesis texts. *Reading and Writing*, 33(2), 239-266.
- [24] Ohta, R., Plakans, L. M., & Gebril, A. (2018). Integrated writing scores based on holistic and



multi-trait scales: A generalizability analysis.  
Assessing writing, 38, 21-36.

iterative integrated assessment tasks.  
Assessing writing, 40, 14-26

[25] Plakans, L., Liao, J.-T., & Wang, F. (2018).  
Integrated assessment research: Writing-into-  
reading. Language teaching, 51(3), 430.

[26] Plakans, L., Liao, J.-T., & Wang, F. (2019). "I  
should summarize this whole paragraph":  
Shared processes of reading and writing in