A Systematic Literature Review of Authentic Assessment in K-12 ESL/EFL Education

Vikineswari S. Mariappan[1]* , **Kamisah Osman[2]**

[1]Faculty of Education, Universiti Kebangsaan Malaysia (UKM), 43600 Bangi, Selangor, Malaysia.
Email: p113215@siswa.ukm.edu.my
[2]Faculty of Education, Universiti Kebangsaan Malaysia (UKM), 43600 Bangi, Selangor, Malaysia.
Email: kamisah@ukm.edu.my

CORRESPONDING AUTHOR (*):
Vikineswari S.Mariappan
(p113215@siswa.ukm.edu.my)

ABSTRACT

This research offers an extensive review and analysis of the use of authentic assessment in K-12 ESL/EFL education. The study follows the PRISMA 2020 statement guidelines for systematic literature reviews, investigating peer-reviewed articles from the SCOPUS, WOS, and ERIC databases (2013-2022). The samples comprised teachers, students, and K-12 education documents; non-eligible criteria were excluded. Out of numerous potential studies, 19 were deemed fitting for the review. The findings indicate growing research interest in authentic assessment, with one publication per year minimum since 2015, and notable contributions from Indonesia. Most of the studies (74%) adopted qualitative designs and utilized different authentic assessment types, such as role play, oral interviews, projects, and portfolios. Key factors affecting the implementation included teacher attributes (assessment literacy, language proficiency, support, time), student attributes (motivation, reflection, time, ability), resources (assessment tools, textbooks, guidebooks, teaching aids), and environmental aspects (class size). The review provides valuable insights for teachers aiming to apply suitable authentic assessment strategies and underlines the importance of considering these factors in lesson planning. It emphasizes the need for more research on teachers' assessment literacy and their experiences in employing authentic assessment in K-12 ESL/EFL education, which is currently an underexplored area.

Contribution/Originality: This study makes an original contribution by providing a comprehensive review of authentic assessment in K-12 ESL/EFL education and highlighting key implementation factors. It further emphasizes the underexplored domain of teachers' assessment literacy, thereby paving the way for future research in this area.

1. Introduction

The increasing importance of English as a global language has made English language education a priority for students worldwide, particularly for those learning English as a Second Language (ESL) or English as a Foreign Language (EFL) (Crystal, 2003). Assessment plays a crucial role in shaping the teaching and learning process in these contexts (Brown, 2004). Traditional assessments, such as standardized tests, often fail to accurately reflect students' performance and do not assess the full range of essential student outcomes (O'Malley & Pierce, 1996). As a result, authentic assessment has emerged as a viable alternative, offering a more holistic evaluation of students' real-life skills and competencies (Wiggins, 1998).

Authentic assessment refers to the process of evaluating students' knowledge and skills through tasks that resemble real-world situations and challenges (Gulikers et al., 2004). In the context of ESL and EFL, authentic assessment aims to measure students' language proficiency and communicative competence using diverse and realistic assessment methods (O'Malley & Pierce, 1996). Literature has provided in-depth findings about the advantages of authentic assessment, such as increased student engagement, more accurate insights into students' language development, and better support for diverse learners (Hart, 1994; Boon et al., 2017; Wiggins, 1998; Darling-Hammond & Snyder, 2000; Gulikers et al., 2004; Herrington & Herrington, 1998).

While many studies have been conducted on authentic assessment in various fields and perspectives, systematic literature reviews on this topic are limited. Only a few systematic literature reviews on authentic assessment have been found, and none of them specifically focus on the K-12 ESL and EFL context (Sokhanvar et al., 2021; Razmawaty & Othman, 2017; Villarroel et al., 2019). This gap in the literature highlights the need for a systematic review that synthesizes and critically analyses the existing research on authentic assessment in the K-12 ESL and EFL context.

The purpose of this systematic literature review is to explore the existing research on authentic assessment in the K-12 ESL and EFL context. This review focuses on K-12 education instead of tertiary education, as the former is a critical stage for language acquisition and development, laying the foundation for future language learning and academic success (Lightbown & Spada, 2006). Additionally, K-12 education often receives less attention in the literature compared to tertiary education, creating a need for further exploration and analysis (Butler & Lee, 2010).

A systematic review was chosen over other types of reviews, such as scoping reviews, because it provides a comprehensive, transparent, and replicable synthesis of the available research, allowing for a more robust analysis of the evidence (Grant & Booth, 2009). The review is guided by the following research questions:
  i.   What are the patterns of authentic assessment study in the ESL/EFL context of K–12 education?
 ii.   What are the types of authentic assessment practised in the ESL/EFL context of K–12 education?
iii.   What are the factors that influence the implementation of authentic assessment in the ESL/EFL context of K–12 education?

The scope of this review encompasses studies focused on K-12 ESL and EFL students, exploring various types of authentic assessment practices and the factors that influence

their implementation. While this review aims to provide a comprehensive overview of the topic, it is important to acknowledge its limitations, such as potential biases in the selection of studies and the challenges associated with comparing and synthesizing findings across diverse educational contexts and methodologies.

Despite these limitations, this systematic literature review seeks to contribute to the understanding of authentic assessment in the K-12 ESL and EFL context, offering valuable insights and recommendations for educators, researchers, and policymakers. By addressing the research questions outlined above, this review aims to identify patterns in the current literature on authentic assessment in K-12 ESL and EFL education, explore the various types of authentic assessment practices used in these contexts, and examine the factors that influence their implementation.

Through this systematic and comprehensive analysis, this review seeks to provide a foundation for future research and practice in the field of authentic assessment in K-12 ESL and EFL education. Ultimately, the findings of this review can help inform educational stakeholders on how to effectively implement and support authentic assessment practices that foster meaningful language learning and development in K-12 students.

2. Methodology

This systematic literature review was conducted in compliance with the "Preferred Reporting Items for Systematic Reviews and Meta-Analysis" (PRISMA 2020 statement). The PRISMA 2020 statement was designed primarily for systematic reviews of studies that evaluate the effects of health interventions, but the checklist items can also be used to report systematic reviews evaluating other interventions, such as social and educational interventions. Therefore, this systematic literature review is reported according to the checklist presented in PRISMA 2020.

2.1. Eligibility Criteria

Table 1 shows details on the inclusion and exclusion requirements. The inclusion criteria were: peer-reviewed empirical research journal articles on authentic assessment in an ESL/EFL context published in SCOPUS, WOS, and ERIC databases from 2013 to 2022 in English. Sample of the studies were teachers, students, or documents related to K–12 education. Other than the stated criteria were excluded from the review.

Table 1: Eligibility Criteria

Criteria	Inclusion	Exclusion
Date	Papers that were published from 2013 to 2022 were included	Papers that were published before 2013 were excluded
Language	English	Studies other than English
Sample	Teachers, students or documents in K12 education	Papers with participants other than K12
Publication	Peer reviewed empirical research journal articles	Non peer reviewed articles, technical reports, guidelines, research synthesizes, books and book chapters were excluded

Setting	Papers that focused on authentic assessment practices in ESL/EFL for K 12 learners	Papers that focused on authentic assessment practices in other educational settings such as higher education and non ESL/ESL excluded
Database	SCOPUS, WOS and ERIC	Other Databases

2.2. Information Sources and Search Strategy

WOS, ERIC and SCOPUS databases were used to find the studies because they are considered central databases for social science research (Taylor, Dempster & Donnelly, 2003). The process of searching the database was completed in June 2022. The main keywords searched were ("authentic assessment*" OR "authentic evaluation*" OR "authentic test*" OR "authentic task*") AND english OR esl OR efl OR elt OR ell OR l2 OR esol) AND NOT (undergraduate* OR university* OR college* OR tertiary* OR "higher education*"). These keywords were searched in the databases mentioned. The search is limited to journal articles published in English from 2013 to 2022.

2.3. Study Selection

The search results' articles were exported into Microsoft Excel. The list of articles was checked for duplicates and title-relatedness. All duplicates and articles that did not fit the K–12 context for ESL/EFL were rejected. The titles and abstracts of all the chosen articles were scrutinised by a two researchers to exclude any that did not address the research questions. The retrieved full text was subjected to further scrutiny about the articles' eligibility.

2.4. Data Extraction and Data Analysis

Each of the selected papers was read completely, and relevant results were recorded using a data extraction form. The data extraction form consists of the following sections:
• General information: author, publication year, title, country
• Research design: method, sample, research questions
• Results and conclusion: answer to the research questions
The next step after completing the form was using MS Excel to assign codes and categories as a form of analysis to answer the research questions.

3. Result

3.1. Study Selection

The deadline for the search was June 13, 2022. The initial search retrieved 111 papers from three databases. 42 papers were found to be duplicates. Following the removal of duplicates, 69 papers were listed for review purposes. After reviewing the titles and abstracts, 34 papers were excluded due to their irrelevancy in the ESL/EFL context of K–12 education. The remaining 35 studies were sought for retrieval. Due to the researcher's lack of access to the document, two studies were unable to be retrieved. When the remaining 33 studies were scrutinised for eligibility, it was discovered that six of them did not address K–12 education, two did not have an ESL/EFL perspective, and six did not involve empirical research. After evaluating all the studies for eligibility, a total of 19

studies were selected for review. The step-by-step search and selection strategy is shown in Figure 1 using the PRISMA 2020 template for systematic literature review.

Figure 1: Flowchart of the Screening and Selection Procedure
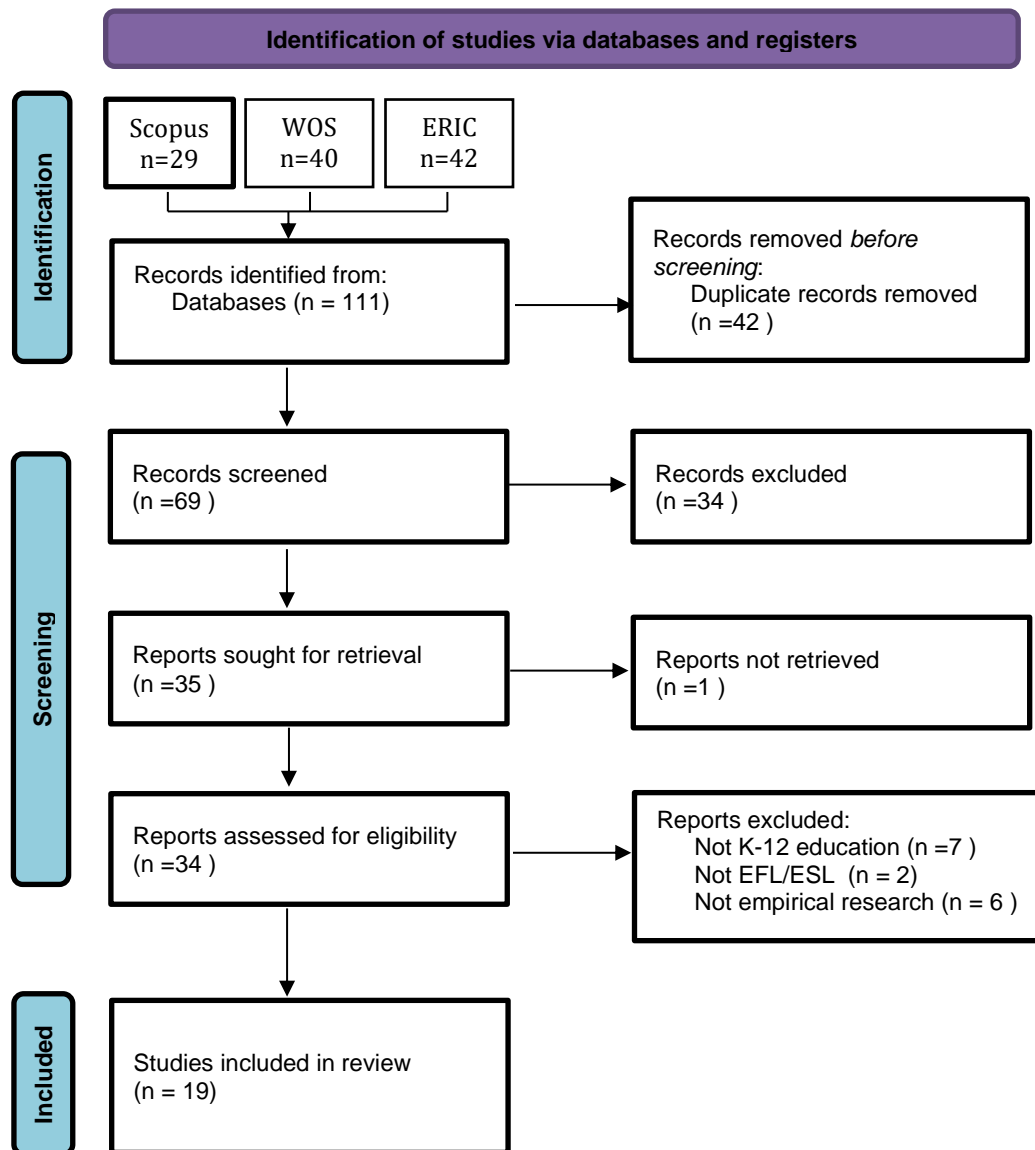
3.2. Study Characteristics

The characteristics of the 19 reviewed papers are presented in Table 2. Table 2 shows various aspects of authentic assessment that have been researched in the context of ESL/EFL for K–12 education. Teacher's perceptions, challenges, effects, implementation, advantages, instruments, and resources have been studied.

Table 2: Characteristics of the Studies

No.	Author & Year	Country	Journal	Method	Sample	Aim
1	Gkogkou and Kofou (2021)	Greece	Languages	Quantitative	120 EFL teachers	Analyze the ways in which Greek EFL instructors respond to communicative testing and their understanding of assessment principles and procedures.
2	Larenas et al. (2021)	Chile	Ensaio	Qualitative	205 instruments 10 teachers 2 university teachers	Examine the fairness, authenticity, validity, reliability, and practicability of 201 assessments made by English teachers.
3	Huang and Jiang (2020)	China	Educational Studies	Mixed Method	-2 tier secondary school(6 teachers from school A and 5 teachers from school B) -30 English teachers from school A and B respectively	examine how Chinese secondary English instructors perceive and use authentic assessment in their classrooms.
4	Muhammad Noor et al. (2020)	Malaysia	International Journal of Evaluation and Research in Education (IJERE)	Qualitative	6 primary excellent teachers	to examine difficulties good teachers encounter when preparing authentic assessment in an ESL environment.

5	Zaim et al. (2020)	Indonesia	International Journal of Instruction	Quantitative	120 secondary EFL	to expose the issues with instructors' on authentic assessment to assess junior high school students' English-speaking abilities, learn what teachers need to know about AA of speaking abilities, and create a model of AA
6	Erwin Rahayu Saputra et al. (2020)	Indonesia	Journal of Education for Teaching	Qualitative	6 English teachers from secondary school	to investigate the adjustments English teachers go through in their assessment during this project
7	Meriem et al. (2019)	Algeria	Jordan Journal of Modern Languages and Literatures	Qualitative	Textbook	To determine the degree to which the CLT tenets are included into the first-stage EFL secondary textbook used in Algeria, "At the Crossroad,"
8	Ratminingsih et al. (2018)	Indonesia	International Journal of Instruction	Quantitative	425 high school students	Analyze how self-evaluation affects pupils' independence and writing ability.
9	Dwi and Lenggahing (2017)	Indonesia	Indonesian Journal of Applied Linguistics	Qualitative	2 English teachers, document and observation	-outlining the use of authentic assessment to gauge students' English-language proficiency - the difficulties in putting into practise - the solutions employed by teachers
10	So et al. (2015)	USA	Educational Technology & Society	Qualitative	20 students	Look at the value of trialogue-based activities, which are computer-mediated conversation-based tasks, as a tool for language assessment
11	Zia et al. (2018)	Iran	Iranian Journal of Language Teaching Research	Qualitative	13 novice and 13 experienced teachers	to investigate the literacy for speaking in the classroom used by new and experienced teachers of English as a foreign language

12	Ortiz and Cuéllar (2018)	Columbia	HOW	Qualitative	8 students from 10th level	to determine how authentic activities affected a group of EFL students' oral production, which was measured in terms of their ability to produce the language in actual communicative situations.
13	Cimen (2022)	Turkey	International Online Journal of Education and Teaching	Qualitative	50 secondary school teachers	to investigate how EFL assessment is used in high school, how it is addressed in the national curriculum (for students in grades 9 through 12), and whether the curriculum and teachers' practices are consistent.
14	Muhammad Noor and Nurahimah (2016)	Malaysia	International Journal of Evaluation and Research in Education	Qualitative	11 primary year 4 children	to talk about how process writing is improved in an English language classroom by using authentic assessment
15	Mohammed and Abdurehman (2020)	Ethiopia	Educational Research and Reviews	Qualitative	24 English teachers from 6 secondary schools	to examine the difficulties in setting up authentic English language exams in Ethiopia's six secondary schools in order to evaluate present difficulties in setting up authentic and communicative language exams.
16	Rastegaran (2016)	Iran	Advances in Language and Literary Studies	Quantitative	62 EFL learners	Examine the impact that authentic tasks have on the acquisition of adjectives.
17	Limberg (2015)	German	Language Teaching Research	Qualitative	3 textbooks	explores the input German textbooks for EFL students give on appologies and what activities and exercises they recommend to help students practise and perform this act in class.
18	Muhammad Noor and	Malaysia	The Malaysian Online Journal of	Qualitative	11 primary year 4 children	investigating the benefits of portfolio assessment for assessing students' writing skills

|						
	Nurahimah (2015)		Educational Science			
19	Karakoc and Bay (2016)	Turkey	Journal of Education and Learning	Qualitative	52 students and 4 teachers	investigating the effectiveness of the authentic task-based approach in the fifth-grade foreign language course at the secondary level

4. Results

4.1. What are the patterns in Authentic Assessment study in ESL/EFL context of K12 education?

The first research question is attempted by answering the distribution of the studies by years, distribution by countries, research methodology used in the selected articles, and samples used in the reviewed articles.

4.1.1. Distribution of the studies by years

As illustrated in Table 2, the descriptive information for RQ 1 has been sourced directly from the articles under review. Meanwhile, Figure 2 showcases the total number of articles on Authentic Assessment and English as a Second Language/English as a Foreign Language (ESL/EFL) for K–12 education that were published each year from 2015 to 2022. From Figure 2, we can conclude that at least one article per year was published from 2015 to 2022 in the context of ESL/EFL for K–12 education. The year 2020 has had five articles published, which is the most compared to other years. It is expected that more articles will be published in the coming years, as authentic assessment is the current trend in assessment.

Figure 2: Total Number of Articles on AA and ESL/EFL for K12 by year
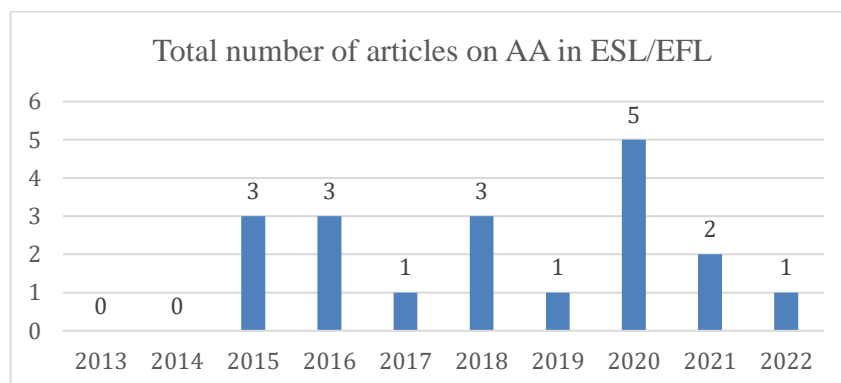

4.1.2. Distribution by Countries

Figure 3, which maps the distribution of articles by country, reveals that authors from Indonesia have contributed the most articles on authentic assessment in the context of ESL/EFL for K–12 education, highlighting Indonesia's significant role in this research area. This is because authentic assessment is one of the major concerns in Indonesia's 2013 curriculum, which promotes students' autonomy and critical thinking (Moch., 2018). Malaysian authors contributed the second most articles, as authentic assessment is the type of assessment practised in classroom-based assessment. In the Classroom Assessment (PBD) handbook, it is stated that PBD is an authentic assessment method without pressuring the students (Kementerian Pendidikan Malaysia, 2019). The Malaysian Education Development Plan (PPPM) 2013–2025 also outlines school assessment, referring to various forms of authentic and contextual assessment that measure the level of mastery of students. (Kementerian Pendidikan Malaysia, 2013).

Turkey and Iran produced two articles each, as the curriculum of these countries also encourages authentic assessment. Other countries such as the USA, Greece, Germany, Ethiopia, Columbia, China, Chile, and Algeria contributed one article each throughout 2013–2022.

Figure 3: Studies by Country
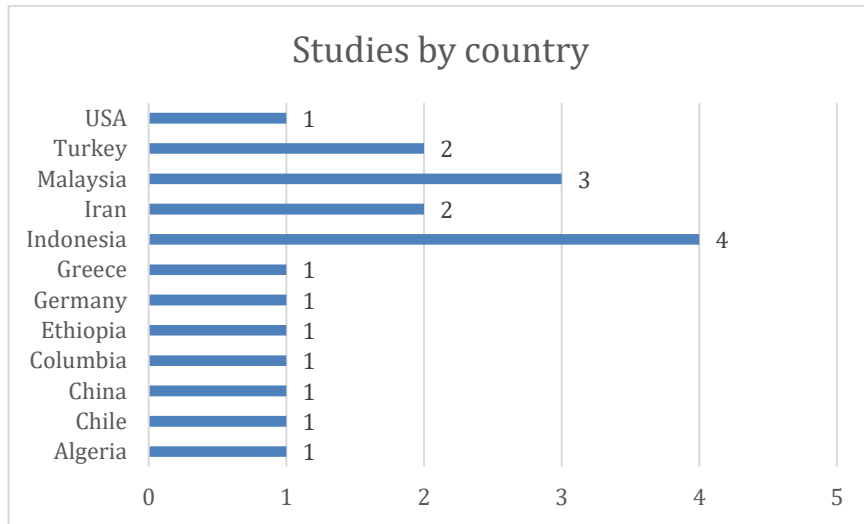
4.1.3. Research Methodology used in the Study

Figure 4, which illustrates the research methods used in the studies, indicates that researchers have utilized three distinct methodologies in conducting research on authentic assessment within an ESL/EFL context for K–12 education. The methods are quantitative methods, qualitative methods, and mixed methods. 74% of the papers reviewed were conducted using the qualitative method, by using document analysis, open ended questionnaire, semi structured interviews, and observation. The quantitative method was used in 21% of the studies reviewed, with questionnaire-based surveys. Only 5% of the papers conducted studies using the mixed method.

Figure 4: Research Method Used in the Study
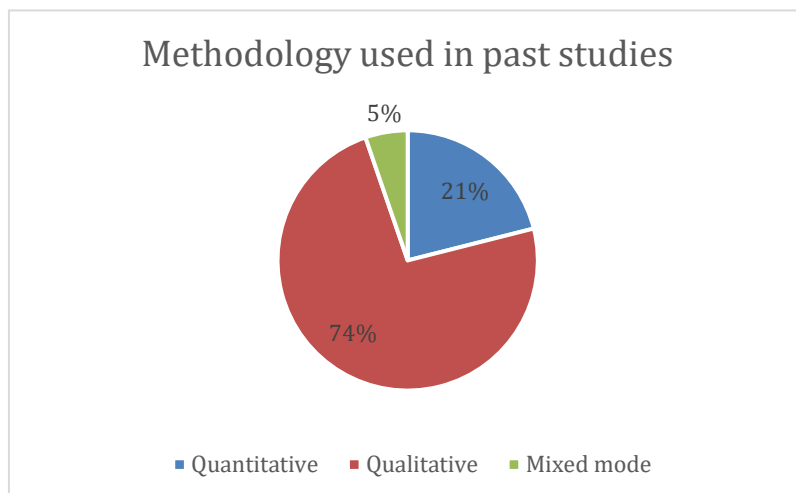
4.1.4. Focus of the Reviewed Study

The study's focus in the publications under review is depicted in Figure 5. It states that seven studies were conducted to examine the results of the use of authentic assessment. Only one study looked at teachers' experience and assessment literacy; two studies looked at resources/instruments and challenges; and two papers looked at teachers' perception and implementation.

Figure 5: Focus of the Study
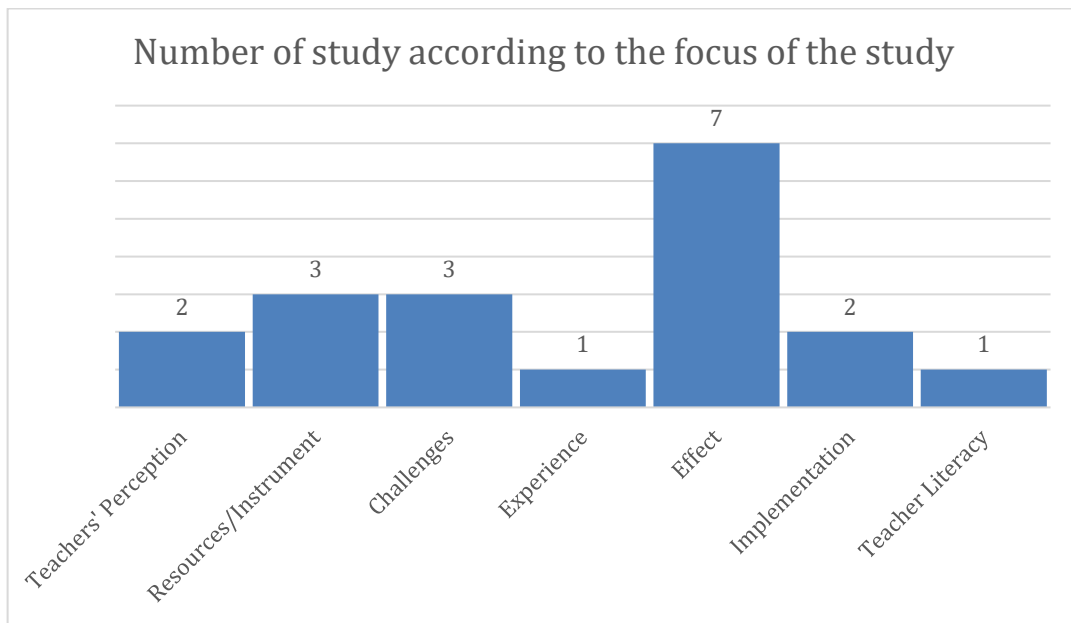
4.1.5. Samples Used in the Reviewed studies

Figure 6 provides a detailed visual representation of the samples used in the reviewed studies, clearly depicting the distribution of samples including teachers, students, documents, and combinations of these. The majority of papers (n = 8) chose teachers as their samples, followed by students (n = 7), documents (n = 2), teacher and student (n = 1), and documents (n = 1).

Figure 6: Types of Study Sample
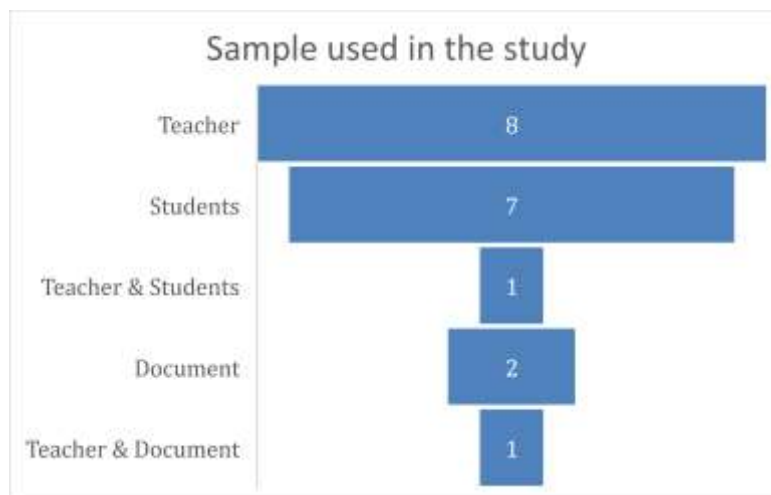
Overall, the findings of research question 1 revealed that from 2013 to mid-2022, there was an increase in the volume of literature on authentic assessment in the context of ESL/EFL for K–12 education. Most of the articles came from countries where English is used as a second language in daily life. As previously mentioned, more articles were from Indonesia because authentic assessment was emphasised in their 2013 curriculum. In terms of methodology, the articles focused mostly on the qualitative method. Qualitative design (74 percent) has a systematic subjective approach to exploring the assessment practices. Quantitative (21 percent) and mixed methods (5 percent) are used in other papers. Finally, most of the samples were taken from teachers, followed by students, documents, teachers and students, and teachers and documents.

4.2. What are the types of authentic assessment practiced in ESL/EFL context of K12 education?

There are various types of authentic assessment used in the classroom. Teachers can choose assessments that serve a specific goal or modify their methods to fit the demands of their students or their lessons (O'Malley & Pierce, 1996). Table 4 and Table 5 show the types of authentic assessment practised in the reviewed studies. All the types of assessments are divided into two categories according to the skills being tested, namely speaking and writing skills.

4.2.1. Authentic Assessment of Speaking Skill

For assessing speaking ability, role play (Zia et al., 2018; Limberg, 2015; Zaim et al., 2020), question and answer or oral interview (Zia et al., 2018; Zaim et al., 2020), monolog (Zia et al., 2018; Ortiz & Cuélla, 2018; Zaim et al., 2020), Picture story (Zia et al., 2018; Zaim et al., 2020; Dwi & Lenggahing, 2017), project (Ortiz & Cuéllar, 2018), pair dialogue (Zaim et al., 2020; Ortiz & Cuéllar, 2018), story or text retelling (Zaim et al., 2020; Dwi & Lenggahing, 2017) and triologue (So et al.,2015) being practised. The types of authentic assessments used to assess speaking skills in the reviewed papers are displayed in Table 3.

Table 3: Types of Authentic Assessment to Assess Speaking Skill

Assessment	f	Articles (author)
Role Play	3	Zia et al. (2018) Limberg (2015) Zaim et al. 2020
Question and Answer (Oral interviews)	2	Zia et al. (2018) Zaim et al. 2020
Monolog	3	Zia et al. (2018) Ortiz and Cuéllar (2018) Zaim et al. 2020
Picture story	3	Zia et al. (2018) Zaim et al. 2020 Dwi and Lenggahing (2017)
Project	1	Ortiz and Cuéllar (2018)

Pair dialogue	1	Zaim et al. 2020 Ortiz and Cuéllar (2018)
Story /text retelling	2	Zaim et al. 2020 Dwi and Lenggahing (2017)
Trialogue	1	So et al. (2015)

Roleplay involves assuming a certain person and setting. Both alone and in groups are acceptable. Limberg (2015) stated that task from the textbook is used for roleplay with the help of a language box that provides additional expression to help students vary their apologetic behaviour when performing it in class. Study by Zia et al. (2018) reveals that more experienced teachers practise role play than novice teachers.

Zia et al. (2018) and Zaim et al. 2020 reported question and answer or oral interview as type of authentic assessment practiced. Oral interview is a method of directly evaluating speaking skills in students (Zaim et al., 2020). In this assessment, the teacher asks questions about a topic and students answer.

According to Ortiz and Cuéllar (2018), speaking tasks that involve sharing personal experiences, discussing likes and hobbies, and planning or creating something tend to be the most helpful ones in facilitating speaking. Monolog or oral reports gives an opportunity for students to practice speaking (Zaim et al.,2020). Zia et al. (2018)stated that monolog is practised by both experienced and novice teachers.

A picture story, also known as a picture-cued description, is a type of authentic assessment in which students describe, tell a story about, provide information about, and express their opinion about a picture. Pictures are ideal to elicit oral production (Zail et al., 2020). Research done by Zaim et al. (2020) shows that picture story is practised only by novice teachers, and experienced teachers have no knowledge about picture story.

Ortiz and Cuéllar (2018) stated that projects can be used to foster oral production among students. In groups, students can organise and present a proposal for a party or create a poster individually about any event and convince their friends to participate. The assignment was enjoyable for the students because it allowed them to make plans for the actual party or event.

Pair dialogue is carried out by separating the students in the class into pairs. The other student must respond to a question with the appropriate pattern after the questioner. The teacher changes pair partners for each round of question and answer (Zaim et al., 2020). Pair dialogue is a type of authentic assessment in which information is exchanged (Ortiz & Cuéllar, 2018).

In the story/text retelling assessment, students read/listen to a story or text. Then, students retell the main ideas or selected details of the text they listened to or read. This kind of evaluation encourages students to recount the events in the story or text (Zaim et al., 2020). The advantage of story/text retelling assessment is that teachers can determine

the reading comprehension, reading strategies, and language level development of the students (O'Malley & Pierce, 1996).

Trialogues are technology-assisted tasks to measure the English language proficiency of the students. Trialogues are interactive virtual conversations between one student and two virtual characters. Each task was designed with conversation patterns between a student and the virtual characters. At the time of the usability study of So et al. (2015), an automated speech recognition (ASR) engine was not incorporated, so it was not able to automatically recognise and respond to students' spoken responses. Therefore, students need to type their responses. Students were also asked to record their responses as speech samples to train an ASR engine at a later stage of development. In the future, trialogue-based tasks can accept spoken input directly, freeing students from typing.

4.2.2. Authentic Assessment of Writing Skills

There are four types of authentic assessment practised in assessing writing skills. There are writing samples (Muhammad Noor & Nurahimah, 2016; Ratminingsih et al., 2018), portfolio (Muhammad Noor & Nurahimah, 2016; Muhammad Noor and Nurahimah, 2015; Huang & Jiang, 2020; Dwi & Lenggahing, 2017); and project (Dwi & Lenggahing, 2017; Huang & Jiang, 2020).

Table 4 shows the types of authentic assessments used to assess writing skills. Writing samples are student generated narrative, expository, persuasive, or reference papers written to meet different purposes (O'Malley & Pierce, 1996). This type of assessment is practised in Muhammad Noor and Nurahimah's (2016) and Ratminingsih et al.'s (2018) studies. Self-assessment plays an important role in writing samples because it gives more opportunity to the students to reflect on their own learning and progress (Ratminingsih et al., 2018). Muhammad Noor and Nurahimah (2016) conducted a writing process that included pair writing, dialogue scripting, picture description, and journal entries to monitor the students' gradational growth in writing abilities.. Motivation in writing among students also increased with the comments and guidance from the teacher (Muhammad Noor & Nurahimah, 2016).

Table 4: Types of Authentic Assessment to Assess Writing Skill

Assessment	f	Articles (author)
Writing sample	2	Muhammad Noor and Nurahimah (2016) Ratminingsih et al. (2018)
Portfolio	3	Muhammad Noor and Nurahimah (2016) Muhammad Noor and Nurahimah (2015) Huang and Jiang (2020) Dwi and Lenggahing (2017)
Project	1	Dwi and Lenggahing (2017) Huang and Jiang (2020)

A portfolio is a deliberate assemblage of students' work that can demonstrate their development over time. Student works chosen by the students themselves or with teacher

help are included in portfolios (O'Malley & Pierce, 1996). Four studies reviewed used the portfolio as an authentic assessment (Muhammad Noor & Nurahimah, 2016; Muhammad Noor & Nurahimah, 2015; Huang & Jiang, 2020; Dwi & Lenggahing, 2017). Findings of the study by Muhammad Noor and Nurahimah (2015) revealed that, from a portfolio, teachers are able to see the progress of their students and provide appropriate guidance and remedial activities to help the students improve. In Muhammad Noor and Nurahimah's (2016) study, the students' writing samples are compiled into a portfolio. The students' writing improved when they became more engaged in using the portfolio.

A project requires students to complete a set of tasks in order to create something (Dwi & Lenggahing, 2017). A project can be conducted individually or in a small group and is often presented through an oral or written report (O'Malley & Pierce, 1996). Project assessment is practised in the studies of Dwi and Lenggahing (2017), Huang and Jiang (2020). Comic production is the project conducted by students in the study of Dwi and Lenggahing (2017); elsewhere in the study of Huang and Jiang (2020), projects in the form of writing letters for real world purposes and responding to news articles have been conducted.

4.3. What are the factors that influence the implementation of authentic assessment in ESL/EFL context of K12 education?

Factors that influence the implementation of authentic assessment in the ESL/EFL context of K-12 education are presented in Table 5. Based on the analysis of the studies presented in Table 5, the factors are divided into four categories: teacher, student, resources, and environment.

Table 5: Factors that Influence the Implementation of Authentic Assessment

Category	f	Articles (Author)
Teacher		
Assessment Literacy	7	Gkogkou and Kofou (2021), Huang and Jiang (2020), Larenas et al. (2021), Zia et al.(2018), Cimen (2022), Muhammad Noor et al. (2020), Erwin Rahayu Saputra et al. (2020)
Language proficiency	2	Mohammed and Abdurehman (2020), Muhammad Noor and Nurahimah (2016)
Support	1	Muhammad Noor et al. (2020)
Time	3	Karakoc and Bay (2016), Muhammad Noor and Nurahimah (2016)
Student		
Motivation	5	Mohammed and Abdurehman (2020), Ortiz and Cuéllar (2018), Muhammad Noor and Nurahimah (2016), Muhammad Noor and Nurahimah (2015), Dwi and Lenggahing (2017)
Reflection	1	Ortiz and Cuéllar (2018)
Time	1	Muhammad Noor and Nurahimah (2016)

Ability	2	Huang and Jiang (2020), Muhammad Noor and Nurahimah (2016)
Resources		
Type of assessment/instrument	5	Larenas et al. (2021), Ortiz and Cuéllar (2018), Zaim et al.(2020), Ratminingsih et al. (2018), So et al. (2015)
Textbook/guidebook/teaching aid	4	Zeghdoud et al. (2019), Limberg (2015), Karakoc and Bay (2016), Mohammed and Abdurehman (2020)
Environment		
Class size	2	Mohammed and Abdurehman (2020) Karakoc and Bay (2016)

Teachers' assessment literacy is the ultimate factor that influences the implementation of authentic assessment in the context of ESL/EFL K–12 education. Teachers are confused about the concept of authentic assessment (Huang & Jiang, 2020; Cimen, 2022). Huang and Jiang (2020) stated that despite having knowledge of authentic assessment, the number of teachers who implemented authentic assessment was rather limited. The teachers were uncertain whether the assessment they designed was authentic or not. This scenario would put learners at risk when not assessed by trained professionals. The study by Zia et al. (2018) also revealed that experienced teachers demonstrated more consistency in assessment literacy compared to novice teachers. Therefore, teachers need training on the implementation of authentic assessment (Muhammad Noor & Nurahimah, 2016) and master the principles of assessment instruments (Larenas et al., 2021), which will enable them to implement authentic assessment efficiently. The professional learning community also increased teachers' awareness and broadened their understanding of assessment literacy and authentic assessment practise (Erwin Rahayu Saputra et al., 2020). Assessment literacy and training can motivate teachers to modify their assessment practises for the benefits of teaching and learning (Gkogkou & Kofou, 2021).

A study by Mohammed and Abdurehman (2020) stated that teachers' language proficiency is one of the main challenges in the implementation of authentic assessment. Pronunciation is not taught to the students because the teachers themselves are not proficient and there are no records of English words in the curriculum. Teachers also need to be clear when giving instructions on how to do a task and explain it clearly to the students by simplifying the instructions. If not, the students would not be able to complete the task (Muhammad Noor & Nurahimah, 2016). Teachers' language proficiency is important for the lesson to be delivered accurately and for the students to accomplish any task given.

According to Muhammad Noor et al. (2020), lack of support from school administration and parents was the greatest challenge in implementing authentic assessment. Teachers had difficulty explaining to parents about authentic assessment as they were only interested in A's in exam papers. Lack of support from the school administration made teachers frustrated and discouraged their motivation. Therefore, support from parents and school administrators is an important factor in the implementation of authentic assessment.

Data collected by Muhammad Noor et al. (2020) proved that more time is needed for process writing. The time allocated is not enough to improve the students' writing skills.

It is a burden for the teachers, as they have to finish the syllabus and keep up with the students' achievements. According to Karakoc and Bay (2016), teachers did not have sufficient time to prepare for authentic assessment. Therefore, time is the factor that influences the implementation of authentic assessment in the EFL/ESL context of K–12 education.

Students' motivation is a significant factor in implementing authentic assessment because language assessment is impossible without students' participation. Teachers had difficulty preparing real life language tests when students don't volunteer to make interactions in the classroom (Mohammed & Abdurehman, 2020). Ortiz and Cuéllar (2018) state that positive effects happen when students are encouraged and supported by their peers and teachers. This is supported by Muhammad Noor and Nurahimah (2016), where students' motivation in writing increased with comments and guidance provided by the teacher. Dwi and Lenggahing (2017) revealed that it was hard to encourage students' to do oral presentations. Teachers encouraged the students to accomplish their task by giving extra points to the student who speaks first in the performance assessment.

Reflection is an important factor because when students' reflect on their performance, they analyse their strengths and weaknesses Ortiz and Cuéllar (2018) Reflection would give room for improvement. Teachers should allow time for reflection when planning assessments.

Time is another factor that influences students' performance in real-world situations. In Muhammad Noor and Nurahimah's (2016) study, students did not have enough time to complete their task, which frustrated them while they were enjoying their writing class. Authentic assessment needs to be planned according to the time frame to avoid disappointment among students.

Huang and Jiang (2020) reported that teachers in their study believed that their students were not mature enough to do authentic assessment as they were not able to take control of their own learning. Students also do not have an idea when it comes to writing assessments, and they lack vocabulary as they have problems constructing sentences (Muhammad Noor & Nurahimah, 2016). Students' inability has an impact on the implementation of authentic assessment and is regarded as one of the factors influencing the implementation of authentic assessment.

Authentic assessment helps to increase students' confidence when communicating because not much emphasis is given to linguistic forms (Ortiz & Cuéllar, 2018). Although language assessment is mainly about communication, teachers still stress grammar and vocabulary knowledge instead of developing the skills of foreign language communication through authentic assessment. Assessment-like tests and quizzes only stress knowledge of language but not its application in communicative situations (Larenas et al., 2021). When self-assessment is practiced, students become aware of their mistakes in grammar with the help of a checklist and rubric. In this manner, students learn to judge their work based on the requirement (Ratminingsih et al., 2018). The requirement is communicated to the students so that they can prepare to achieve the required performance level. This type of assessment is considered a path to students' progress and achievement in the classroom (Zaim et al., 2020). A assisted type of assessment like trialogue is able to engage students in an authentic experience and is perceived positively by students (So et al., 2015). Therefore, the types of authentic assessment and instrumentation need to be selected accordingly to obtain the desired outcome from the students.

According to Limberg (2015), a textbook is an important resource that offers a variety of opportunities to receive and practise pragmatic competence in specific areas of language use. Teachers in certain areas depend solely on textbooks to prepare their authentic assessments, as there were no authentic reading materials available in the school library (Mohammed & Abdurehman, 2020). Some textbooks offer authentic tasks, sources, pair work, and picture descriptions. These activities give students opportunities to acquire various real life skills and social skills (Zeghdoud et al., 2019). Guidebooks are also important, as it is troublesome to design activities if there is no guidance activity and sufficient information provided in the guidebook (Karakoc & Bay, 2016). Therefore, textbooks should be properly equipped with various communicative principles (Zeghdoud et al., 2019). Besides that, a teaching aid like a tape recorder is vital equipment because it can be used to imitate native pronunciation. The absence of a tape recorder made assessment difficult as the teachers were not able to pronounce words like native speakers (Mohammed & Abdurehman, 2020).

According to Mohammed and Abdurehman (2020) and Karakoc and Bay (2016), the teacher's ability to implement authentic language testing is hampered by large class sizes. Assessing students' performance and peer-to-peer communication in a large class is difficult. Academic dishonesty also occurs due to the large class because the teacher is not able to control the misbehaving student during the assessment. The nature of authentic assessment needs a more interactive environment, so classroom size is a factor that influences the implementation of authentic assessment in the ESL/EFL context of K–12 education.

5. Discussion

This systematic literature review aimed to explore the existing research on authentic assessment in the K-12 ESL and EFL context by addressing three main research questions. The findings provide valuable insights into the patterns of authentic assessment research, the types of authentic assessment practiced, and the factors that influence their implementation in the ESL/EFL context of K-12 education.

The upward trend in research interest on authentic assessment reflects its growing importance in contemporary educational practices, particularly in the context of second language acquisition. The prominence of authentic assessment in countries where English is used as a second language highlights the global shift towards more authentic and contextual assessment practices. The prevalence of qualitative research methodologies, while providing in-depth and context-specific insights, suggests potential for further investigation using quantitative and mixed methods approaches. Moreover, the focus on the outcomes of authentic assessment indicates the need for research on other aspects, such as teachers' experiences, assessment literacy, resources, instruments, and challenges associated with its implementation.

The various types of authentic assessment methods identified in the review, particularly those targeting speaking and writing skills, underscore the importance of assessing language abilities in a more realistic and meaningful manner. These assessments not only provide valuable insights into students' language abilities but also foster their engagement and motivation in the learning process. Further research could explore the use of authentic assessments for other language skills, such as reading and listening, as well as investigate the factors that influence teachers' choice of assessment methods.

The factors influencing the implementation of authentic assessment can be categorized into teacher-related factors, student-related factors, resources, and environment. Addressing these factors through teacher training, student motivation, appropriate resources, and conducive environments can enhance the effectiveness of authentic assessment and ultimately lead to improved language learning outcomes for students. It is essential for educational stakeholders to consider these factors when implementing and supporting authentic assessment practices in K-12 ESL and EFL classrooms.

The findings of this systematic literature review contribute to the understanding of authentic assessment in the K-12 ESL and EFL context and offer valuable insights and recommendations for educators, researchers, and policymakers. By synthesizing and critically analyzing the existing research, this review provides a foundation for future research and practice in the field of authentic assessment in K-12 ESL and EFL education. It is our hope that the insights gained from this review can help inform educational stakeholders on how to effectively implement and support authentic assessment practices that foster meaningful language learning and development in K-12 students.

6. Implication and Recommendation

Based on the findings of this systematic literature review, several implications and recommendations can be drawn for educators, researchers, and policymakers in the field of K-12 ESL and EFL education.

The study highlights the importance of assessment literacy and language proficiency among teachers for the successful implementation of authentic assessment. It is recommended that educational institutions provide ongoing professional development and training programs for teachers to ensure they are well-equipped with the necessary knowledge and skills for implementing authentic assessment. These programs should also address the language proficiency of teachers, as this directly impacts their ability to design and deliver assessments effectively.

To enhance the effectiveness of authentic assessment, it is crucial for educators to plan assessments that foster student motivation and reflection. Teachers should develop assessment tasks that engage students in meaningful and authentic language use, as well as provide opportunities for self-assessment and reflection on their progress. This will not only help students to better understand their strengths and weaknesses but also contribute to their overall language development.

Another important consideration is the availability and quality of resources, such as textbooks, guidebooks, and teaching aids, that support the implementation of authentic assessment. Policymakers and curriculum developers should ensure that these resources incorporate communicative principles and provide valuable guidance for teachers in designing authentic assessments. Additionally, investments in teaching aids, like tape recorders or other digital tools, can facilitate the assessment process and enhance students' language learning experiences.

Finally, the study findings emphasize the role of the learning environment, particularly class size, in the implementation of authentic assessment. Educational institutions should consider class size when planning for the implementation of authentic assessment in the ESL/EFL context of K-12 education. Smaller class sizes can facilitate more effective

assessment by enabling teachers to monitor students' performance more closely and foster better peer-to-peer communication.

In conclusion, this systematic literature review offers valuable insights and recommendations for improving the implementation of authentic assessment in K-12 ESL and EFL education. By addressing the factors identified in this study, educational stakeholders can work towards creating more effective and meaningful assessment practices that ultimately enhance language learning outcomes for students.

7. Conclusion

The systematic literature review of authentic assessment in the K-12 ESL and EFL context provides valuable insights that contribute to a deeper understanding of this important pedagogical tool. The upward trend in the research interest in this field underscores the growing recognition of the potential of authentic assessments in language learning.

Indonesia, having contributed the most articles in this sphere, indicates a regional interest in developing and perfecting these assessment methods. This study has revealed that qualitative research dominates the field, reflecting the complexity and the multidimensional nature of authentic assessments that are best explored through qualitative methods.

Various types of authentic assessments have been practiced, with an emphasis on speaking and writing skills, demonstrating the flexible and versatile nature of these assessments. Role plays, oral interviews, monologs, picture stories, projects, pair dialogues, story retellings, trialogues, writing samples, and portfolios all provide diverse, engaging, and effective ways of assessing student learning.

Influencing factors in the implementation of authentic assessment were categorized into teacher-related, student-related, resources, and environmental factors. Teachers' assessment literacy and language proficiency, alongside the support and time they have, were crucial for successful implementation. Students' motivation, reflection, time, and ability also played a significant role. The types of assessment instruments, textbooks, guidebooks, and teaching aids available, as well as the class size, were other essential considerations.

In conclusion, this systematic literature review underlines the importance and complexity of implementing authentic assessment in the K-12 ESL and EFL context. It highlights the need for teachers, policymakers, and researchers to consider a variety of factors to ensure successful implementation and to enhance the learning outcomes of students. Further research should focus on areas such as improving teachers' assessment literacy and exploring their experiences in implementing authentic assessment. These areas are currently underexplored, yet they significantly influence the successful implementation of authentic assessments.

Funding

This study received no funding.

Conflict of Interest

The authors declare that there are no conflicts of interest regarding the research, authorship, or publication of this article, and confirm that no potential conflicts exist.

References

Boon P. Y., Lee L. H., & Aeria, L. A. (2017). *Pentaksiran dalam pendidikan*. Shah Alam: Penerbit Oxford Fajar.

Brown, H. D. (2004). *Language assessment: principles and classroom practices*. Pearson Education.

Butler, Y. G., & Lee, J. (2010). The effects of self-assessment among young learners of English. *Language Testing, 27*(5), 671–689.

Cimen,S.S. (2022). Exploring EFL assessment in Turkey:Curriculum and teacher practices. *International Online Journal of Education and Teaching*, *9*(1) , 531-550.

Crystal, D. (2003). *English as a global language*. Cambridge University Press.

Darling-Hammond, L., & Snyder, J. (2000). Authentic assessment of teaching in context. *Teaching and Teacher Education,* *16*(5–6), 523–545. https://doi.org/10.1016/S0742-051X(00)00016-9

Dwi Rukmini & Lenggahing Asri Dewi Eko Saputri. (2016). The authentic assessmeny to measure students' English productive skills based on 2013 curriculum. *Indonesian Journal of Applied Linguistics*, *7*(2), 263-273. https://doi.org/10.17509/ijal.v7i2.8128

Erwin Rahayu Saputra, Fuad Abdul Hamied & Didi Suherdi (2020).The development of beliefs and practices of language assessment literacy:does a professional learning community help? *Journal of Education for Teaching, 46*(3), 414-416. https://doi.org/10.1080/02607476.2020.1761250

Gkogkou, E., & Kofou, I. (2021). A toolkit for the investigation of Greek EFL teachers' assessment literacy, *Languages, 6*(4), 188. https://doi.org/10.3390/languages6040188

Grant, M.J., & Booth, A. (2009). A typology of reviews: an analysis of 14 review types and associated methodologies. *Health Information & Libraries Journal, 26*(2), 91-108.

Gulikers, J. T. M., Bastiaens, T. J., & Kirschner, P. A. (2004). A five-dimensional framework for authentic assessment. *Educational Technology Research and Development, 52*, 67–86. https://doi.org/10.1007/BF02504676

Hart, D. (1994). *Authentic assessment:A handbook for educators*. United States of America: Edison-Wesley.

Herrington, J., & Herrington, A. (1998). Authentic assessment and multimedia: how university students respond to a model of authentic assessment. *Higher Education Research and Development, 17*(3), 305-322.

Huang, R., & Jiang, L. (2021). Authentic assessment in Chinese secondary English classrooms: teachers' perception and practice. *Educational Studies*, *47*(6), 633-646. https://doi.org/10.1080/03055698.2020.1719387

Karakoc, B., & Bay, E. (2016). Opinions of the teachers related with the positive and negative aspects of the authentic task based approach in foreign language education. *Journal of Education and Learning, 5*(3), 166-172.

Kementerian Pendidikan Malaysia (2013). *Pelan Pembangunan Pendidikan Malaysia 2013 – 2025: Pendidikan prasekolah hingga lepas menegah*. Putrajaya.

Kementerian Pendidikan Malaysia (2019). *Panduan Pelaksanaan Pentaksiran Bilik Darjah Edisi ke -2*.Putrajaya.

Larenas, C. D., Diaz, A. J., Orellana, Y. R., & Villalon, M. J. S. (2021). Exploring the principles of English assessment instruments, *Ensaio*, *29*(111), 461-483. https://doi.org/10.1590/S0104-403620210002902851

Lightbown, P. M., & Spada, N. (2006). *How Languages Are Learned (3rd ed.)*. Oxford: Oxford University Press.

Limberg, H. (2016). Teaching how to apologize:EFL teaxtbook and pragmatic input. *Language Teaching Research*, *20*(6), 700-718. https://doi.org /10.1177/1362168815590695

Meriem Zeghdoud, Fawwaz Al-Abed Al-Haq, & Dina Al-Jamal (2019). Communicative English Language Teaching Principles in Action: Moving from Theory to Examples. *Jordan Journal of Modern Languages and Literatures, 11*(2), 225-246.

Moch. Said Mardjuki (2018). English teachers' perception on the use of authentic assessment in 2013 curriculum. *Indonesian Journal of English Language Teaching and Applied Linguistics*, *2*(2), 151-163.

Mohammed Shure Wado & Abdurehman Osman Seid (2020). Exploring challenges of preparing authentic English language tests in secondary and preparatory schools: The case of Gamo, Gofa and Konso zones in Southern Ethiopia. *Educational Research and Reviews*, *15*(8), 433-437. https://doi.org/10.5897/ERR2020.3965

Muhammad Noor Abdul Aziz & Nurahimah Mohd Yussof (2015). Using portfolio to assess rural young learners' writing skills in English Language classroom. *The Malaysian Online Journal of Educational Science*, *3*(4), 46-54.

Muhammad Noor Abdul Aziz & Nurahimah Mohd Yussof (2016).Improving process writing with the use authentic assessment. *International Journal of Evaluation and Research in Education (IJERE)*, *5*(3), 200-204.

Muhammad Noor Abdul Aziz, Nurahimah Mohd Yusoff & Mohd Faiz Mohd Yaakob (2020). Challenges in using authentic assessment in 21st century ESL classrooms. *International Journal of Evaluation and Research in Education*, *9*(3), 759-768. https://doi.org.10.11591/ijere.v9i3.20546

O'Malley, J. M., & Pierce, L. V. (1996). *Authentic assessment for English Language learners. Practical approaches for teachers*. United States of America: Longman

Ortiz, S. M. R., & Cuéllar, M. T. A. (2018). Authentic tasks to foster oral production among English as a foreign language learners. *HOW*, *5*(1), 51-68.

Rastegaran, S. (2016). The effect of using authentic task on teaching adjectives. *Advances in Language and Literacy Studies*, *7*(2), 18-24. http://dx.doi.org/10.7575/aiac.alls.v.7n.2p.18

Ratminingsih, N. M., Marhaeni, A. A. I. N., & Vigayanti, L. P. D. (2018). Self assessment: The effect on students' independence and writing competence. *International Journal of Instruction*, *11*(3), 277-290. https://doi.org/10.12973/iji.2018.11320a

Razmawaty Mohamed & Othman Lebar. (2017). Authentic assessment in assessing higher order thinking skills. *International Journal of Academic Research in Business and Social Sciences*, *7*(2), 466-476.

So, Y., Zapata-Rivera, D., Cho, Y., Luce, C., & Battistini, L. (2015). Using trialogues to measure English Language skills. *Educational Technology & Society, 18*(2), 21-32.

Sokhanvar, Z., Salehi, K., & Sokhanvar, F. (2021). Advantages of authentic assessment for improving the learning experience and employability skills of higher education students: A systematic literature review. *Studies in Educational Evaluation, 70*(101030). https://doi.org/10.1016/j.stueduc.2021.101030

Taylor, B. J., Dempster, M., & Donnelly, M. (2003). Hidden gems: Systematically searching electronic databases for research publications for social work and social care. *British Journal of Social Work*, *33*(4), 423-439.

Villarroel, V., Boud, D., Bloxham, S., Bruna, D., & Bruna, C., (2020) Using principles of authentic assessment to redesign written examinations and tests, *Innovations in Education and Teaching International, 57*(1), 38-49, https://doi.org/10.1080/14703297.2018.1564882

Wiggins, G.P. (1998). *Educative assessment: Designing assessments to inform and improve student performance*. San Francisco, CA: Jossey-Bass.

Zaim, M., Refnaldi & Safnil Arsyad (2020). Authentic assessment for speaking skills: Problem and solution for English secondary school teachers in Indonesia. *International Journal of Instruction, 13(*3), 587-604. https://doi.org/10.29333/iji.2020.13340a

Zeghdoud, M., Al-Abed Al-Haq, F., & Al-Jamal, D. (2019). Communicative English language teaching principles in action: Moving from theory to examples. *Jordan Journal of Modern Languages and Literatures, 11*(2), 225-246.

Zia Tajeddin, Minoo Alemi & Hasti Yasaei (2018). Classroom assessment literacy for speaking: Exploring novice and experienced English language teachers' knowledge and practice. *Iranian Journal of Language Teaching Research*, *6*(3), 57-77.