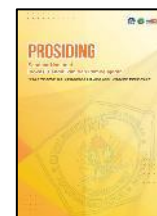




*Prosiding*

Seminar Nasional Inovasi pendidikan dan Pembelajaran  
Fakultas Pendidikan Bahasa dan Seni  
IKIP PGRI Bojonegoro

Tema "Inovasi pendidikan dan Pembelajaran di era digital untuk Pengalaman Belajar  
Imersif"



## The Practices of English Performance Assessment in CLIL Classroom Setting

Rosdina Asyidda Elbirka<sup>1</sup>, Ima Isnaini Taufiqur Rohmah<sup>2</sup>, Ayu Fitriyaningsih<sup>3</sup>

English Education Department, IKIP PGRI Bojonegoro, Indonesia

[rossdinaelbirka@gmail.com](mailto:rossdinaelbirka@gmail.com) , [isnainiima@ikippgribojonegoro.ac.id](mailto:isnainiima@ikippgribojonegoro.ac.id) ,

[ayu\\_fitrianingsih@ikippgribojonegoro.ac.id](mailto:ayu_fitrianingsih@ikippgribojonegoro.ac.id)

**abstrak** – Content and Language Integrated Learning (CLIL) merupakan pendekatan pembelajaran yang menggabungkan penguasaan materi pelajaran dengan pembelajaran bahasa asing, khususnya dalam konteks sekolah dasar. Di kelas CLIL, guru dituntut untuk menilai tidak hanya pemahaman siswa terhadap materi pelajaran, tetapi juga kemampuan berbahasa Inggris mereka. Namun, penelitian mengenai praktik penilaian performa bahasa Inggris dalam konteks CLIL di sekolah dasar Indonesia masih terbatas. Penelitian ini bertujuan untuk mendeskripsikan praktik penilaian performa bahasa Inggris serta mengidentifikasi hambatan yang dihadapi guru dalam proses penilaian tersebut. Penelitian ini dilakukan di salah satu sekolah dasar berbasis CLIL di Bojonegoro dengan menggunakan pendekatan studi kasus kualitatif. Data dikumpulkan melalui observasi kelas, wawancara, kuesioner, dan dokumentasi. Hasil penelitian menunjukkan bahwa: (1) penilaian performa bahasa Inggris dilakukan melalui asesmen terintegrasi, terutama dengan menggunakan tes tertulis dan observasi guru selama kegiatan pembelajaran; dan (2) guru menghadapi beberapa hambatan, termasuk keterbatasan instrumen penilaian khusus serta minimnya strategi yang sesuai untuk mengevaluasi performa bahasa Inggris siswa dalam kerangka CLIL. Temuan ini memberikan gambaran mengenai praktik nyata dan tantangan dalam penilaian performa bahasa Inggris di kelas berbasis CLIL. Studi ini menyimpulkan perlunya alat penilaian yang lebih terstruktur dan pelatihan bagi guru untuk mendukung evaluasi kemampuan bahasa Inggris secara efektif dalam lingkungan CLIL.

**Kata kunci** – CLIL, Kecakapan Bahasa Inggris, Penilaian, Sekolah dasar

**Abstract** – Content and Language Integrated Learning (CLIL) is an educational approach that combines subject content learning with foreign language acquisition, particularly in primary school settings. In CLIL classrooms, teachers are expected to assess not only subject knowledge but also students' English performance. However, research on how English performance is assessed in Indonesian primary CLIL contexts remains limited. This study aims to describe the practices of English performance assessment and to identify the barriers teachers face in conducting such assessments. Conducted in one CLIL-based elementary school in Bojonegoro, this research employed a qualitative case study approach with data collected through classroom observations, interviews, questionnaire, and documentation. The findings reveal that: (1) English performance was assessed through integrated assessments, primarily using written tests and teacher observations during classroom activities; and (2) teachers encountered several barriers, including a lack of specific assessment instruments and limited

strategies tailored to evaluate students' English performance within a CLIL framework. These findings provide insights into the real practices and challenges of English performance assessment in CLIL-based classrooms. The study concludes the need for more structured assessment tools and teacher training to effectively support English performance evaluation in CLIL environments.

**Keywords** – CLIL, English Proficiency, Assessment, Elementary School.

## INTRODUCTION

Content and Language Integrated Learning (CLIL) is an educational innovation that combines content education with language acquisition. The primary aim of CLIL is to enhance both areas simultaneously by integrating subject matter teaching with foreign language learning. This approach is recognized for its potential to increase student motivation, improve language proficiency, promote cognitive development, and foster cross-cultural awareness. As a result, CLIL has gained widespread attention and recognition globally, particularly in the context of second language learning (Eurydice, 2006).

The concept of CLIL originated in Europe during the 1990s and was introduced by experts from educational administration, research, and practice. It quickly gained popularity across Europe and has since expanded beyond the region to countries around the world, including in Asia. Researchers have identified CLIL as a powerful instructional strategy for enhancing language learning, specifically in second or foreign language classrooms (Ball & Lindsay, 2010; Hüttner & Rieder-Bunemann, 2010). Its approach fosters both content and language learning in an integrated manner, benefiting students' overall academic and linguistic development.

In Indonesia, English has traditionally been taught as a separate subject. However, at certain primary schools, English is increasingly being incorporated as a medium of instruction in various subjects like Mathematics, Science, Social Studies, and History, aligning with the principles of CLIL. This shift requires adjustments in the curriculum, as schools implement a combination of the Indonesian National Curriculum and international frameworks to support bilingual education. As part of this approach, schools often introduce International Class Programs (ICP), which offer a bilingual curriculum to students (Puspitasari et al., 2016; Rachmajanti et al., 2015).

With the integration of English into subject areas such as Science, evaluating students' English performance has become crucial. A preliminary study revealed that parents, as key stakeholders, are particularly interested in receiving detailed reports about their children's English language development. Thus, understanding how teachers assess English performance in CLIL classrooms is important for improving teaching practices and ensuring that students' language proficiency is accurately measured and reported.

Assessment in CLIL classrooms is learner-centered, with teachers evaluating students' progress based on specific criteria that reflect their ability to demonstrate practical language skills. These assessments often take place at the end of a course or unit, involving the compilation and organization of student work as evidence of progress. In this context, assessment serves not only as a means of measuring proficiency but also as a tool for providing valuable feedback to both students and teachers to improve the learning process (Ross, S. 2005).

Ross, S. (2005) categorizes assessment into two types: language assessment as quantification and assessment for educational purposes. The first type measures the level of proficiency a student has achieved in acquiring specific language skills, while the second type focuses on integrating assessment strategies into instruction to enhance learning outcomes. In the CLIL context, assessment is seen as an integral component of both teaching and learning, aimed at improving students' language skills and content knowledge simultaneously.

However, while many studies have explored the implementation and benefits of CLIL in various contexts, including Indonesia, there is a noticeable lack of research focusing specifically on how English performance is assessed in CLIL classrooms. Previous studies have highlighted the inadequacy of assessment practices in evaluating students' English performance within these settings. For instance, Rohmah et al. (2021) found that current assessment practices in Indonesian CLIL classrooms are inadequate for evaluating the full range of students' English language skills. Despite these findings, little attention has been paid to the specific assessment methods employed by teachers and the challenges they encounter in bilingual classroom settings such as International Class Programs.

Given the growing implementation of CLIL in various educational settings, particularly in International Classroom Programs in Indonesia, it is essential to explore how teachers assess English performance within this framework. Understanding teachers' assessment practices will provide valuable insights into how English proficiency is evaluated in a dual-focus learning environment. Moreover, investigating these practices can help identify challenges and barriers that educators face when assessing English performance.

The primary focus of this study is to examine how teachers conduct assessments of English performance in CLIL classrooms. The researchers also aims to investigate the challenges that teachers face when conducting these assessments. By understanding these practices and barriers, this study will contribute to improving assessment strategies in CLIL classrooms, supporting teachers in their efforts to effectively evaluate students' English proficiency.

The objectives of this study are twofold: first, to describe the assessment practices that teachers use in CLIL classrooms, and second, to identify the challenges they encounter when assessing English performance. By addressing these objectives,

this study will provide insights that can help refine current practices and overcome obstacles that hinder effective assessment.

This study is significant from three perspectives. Theoretically, it contributes to the existing body of knowledge on English performance assessment in CLIL classrooms, shedding light on the barriers teachers face. Pedagogically, the research supports teachers by providing practical insights into common assessment practices and offering guidance on addressing challenges. Practically, the findings offer actionable solutions for teachers to improve their assessment methods, ensuring more accurate and effective evaluation of students' English performance in CLIL classrooms.

## METHOD

This study employed a qualitative approach using a single case study design to investigate the English performance assessment practices in a CLIL (Content and Language Integrated Learning) classroom. Referring to Creswell (2007), a case study allows for an in-depth analysis of a bounded system, focusing on a specific context and participant group. In this study, the case was bounded to one classroom, examining how English performance was assessed through CLIL instruction. The case study approach was selected due to its suitability in exploring complex phenomena within their real-life context, especially when the boundaries between the phenomenon and context are not clearly defined.

The use of a holistic single case study design enabled the researchers to view the classroom and its assessment practices as a whole, without dividing the focus into smaller sub-units. This approach aligns with Yin's (2003) perspective that a single case study can offer rich insights into a unique or critical case. Although holistic in nature, the study also incorporated elements of an embedded case study by examining various forms of assessments and their implementation. This combination provided a broader and deeper understanding of how English performance assessments were integrated within the CLIL setting.

This study took place at SDI Luqman Al Hakim, a private Islamic elementary school located in Bojonegoro, East Java. Known for its integration of Islamic values and CLIL instruction, the school provided an ideal setting to study how English was taught and assessed within content subjects. English was not taught as a separate subject but embedded within content areas such as science, allowing students to simultaneously learn language and subject matter. The CLIL approach at the school created a unique bilingual learning environment that was particularly relevant to the aims of the study.

The data collection was conducted over the course of February to March 2025, during the second semester of the 2024-2025 academic year. This time frame allowed the researchers to observe naturally occurring classroom activities and assessment

events. Observations were scheduled based on the regular timetable, focusing on lessons where English and content learning were integrated. These sessions provided valuable opportunities to witness how teachers carried out performance assessments in real-time, capturing authentic classroom interactions and assessment practices.

In this study, the researchers focused on one teacher and one co-teacher involved in delivering science content using English as the medium of instruction. The selection of participants was based on their direct involvement in CLIL instruction and assessment, ensuring the data collected would reflect relevant practices. In addition to the teachers, the students in the observed class also became an important source of contextual data, particularly regarding their responses to the assessments conducted in the classroom.

To collect the data, the researchers used four main techniques: observation, interview, documentation, and questionnaire. Classroom observations were conducted using a non-participant approach, meaning the researchers did not intervene in the teaching process but observed assessment activities as they unfolded. These observations were aimed at identifying how performance assessment was carried out, including the tools used and the teacher-student interactions during assessments. The observation also extended to student responses and engagement during tasks.

Semi-structured interviews were conducted with the teacher and co-teacher to obtain more detailed information about their assessment strategies, the rationale behind them, and the challenges they faced. These interviews provided the opportunity for the researchers to explore the teachers' perspectives more deeply. Occasionally, informal interviews were also used to clarify certain points observed during classroom activities.

Documentation analysis was carried out to verify and support the findings obtained from observations and interviews. The reviewed documents included lesson plans, assessment instruments, and school policy documents related to language assessment. These documents helped to reveal whether the stated practices aligned with the implemented ones, thus enhancing the credibility of the data. Document analysis served as a complementary method to understand the structure and criteria used in assessing students' English performance.

Additionally, questionnaires were distributed to the teachers to identify specific challenges or barriers they experienced in conducting English performance assessments within the CLIL framework. The use of questionnaires allowed the researchers to gather more structured responses regarding difficulties in planning, implementing, or evaluating assessments, thus enriching the data collected from interviews and document analysis.

Data were analyzed using qualitative methods as outlined by Miles and Huberman (1994), consisting of data reduction, data display, and conclusion drawing. In the data reduction stage, the researchers filtered the raw data by selecting and organizing information relevant to the study objectives. The data were then displayed using narrative descriptions and tables to facilitate the identification of patterns and themes. Finally, conclusions were drawn based on the patterns observed, and verification was done through triangulation to ensure that the findings were valid and consistent across different sources.

To validate the findings, this study employed triangulation by combining multiple data sources and collection methods. Method triangulation was applied by using interviews, observations, questionnaire and document analysis. Meanwhile, source triangulation involved collecting data from both teachers and students. This approach followed Bryman's (2006) and Leech & Onwuegbuzie's (2007) suggestions on ensuring trustworthiness in qualitative research. Triangulation increased the credibility and dependability of the study by confirming that the findings were not dependent on a single source or technique.

## FINDING AND DISCUSSION

### Findings

In the context of Content and Language Integrated Learning (CLIL), language performance assessment becomes one of the important components to ensure that students not only understand the lesson content but also develop in foreign language skills, especially English. However, in practice, assessment in CLIL classrooms is not always done systematically, especially when teachers do not come from a language education background. This study examines how science teachers in elementary schools implement English performance assessment in CLIL learning and identifies the barriers they face during the process. The findings cover two main focuses, namely the form of assessment practices implemented and the various obstacles that arise in its implementation.

1. English performance assessment practiced by Science teachers in CLIL classroom

Based on the study findings, English performance assessment conducted by science teachers in the context of CLIL (Content and Language Integrated Learning) classes is conducted in an integrated way with lesson content assessment. This assessment is implemented through two main forms, namely written test and teacher observation. The written test is conducted at the end of the semester and is prepared by the Teachers Group Discussion forum with reference to the curriculum standards. The questions are presented in English to measure students' understanding of the science material as well as their

English language skills. The structure of the written test consists of a student identity section and questions covering the material for one semester, such as answering true-false, selecting pictures, and marking objects based on the concepts learned, for example on how seeds spread. The test results are then analyzed by teachers and discussed in teacher meetings to formulate student progress reports, which include not only their mastery of the science material, but also their English performance. This approach reflects the CLIL principle that emphasizes the simultaneous integration of language and content learning.

In addition to written tests, science teachers also assess English performance through observation during the learning process. This observation is conducted informally and continuously, with the aim of seeing how students use English in daily interactions in the classroom. Teachers pay attention to students' use of English in answering questions, discussing, working in groups, as well as in oral presentation tasks. This observation does not use a special assessment instrument, but is based on direct observation of students' attitudes and communication skills. The results of this observation are taken into consideration by the teacher in preparing the learning outcome report, especially in assessing the aspects of speaking skills and students' active involvement in CLIL-based learning. This approach is in line with the characteristics of CLIL learning which emphasizes the natural use of the target language in the context of the lesson content.

## 2. The barriers faced by science teachers related to the practice of English performance assessment in CLIL classroom

The primary barrier in English language assessment in CLIL (Content and Language Integrated Learning) classes is the Lack of language assessment tools. Teachers struggle because they do not have adequate tools to assess students' language abilities. Without clear assessment criteria and specific testing instruments, the assessment process becomes ineffective in measuring students' language skills.

Another issue is the Lack of the development of criteria for assessing the language skills. Many teachers have not developed specific criteria based on language learning objectives. The existing criteria are often vague and mixed with content assessment, making it difficult to evaluate language skills such as pronunciation, vocabulary, grammar, or students' ability to express themselves.

Additionally, there is a Lack of attention of the instruments to assess the language skills. Assessment instruments, such as tests, rubrics, and observation sheets, are not specifically designed to measure listening,

speaking, reading, or writing skills in English. As a result, assessments become incomplete and fail to address the language aspects that should be evaluated.

In addition to the lack of assessment tools, there is also a Lack of language assessment strategies. Many teachers do not use appropriate approaches to assess English language proficiency within the CLIL context. Without systematic strategies, such as formative language assessments or the use of language portfolios, assessments often focus solely on content rather than language skills.

In conclusion, these barriers highlight the need for more focused language assessment tools and strategies. To improve assessment effectiveness in CLIL classes, teachers must have specific assessment instruments, clear criteria, and structured approaches to assess students' language proficiency.

## Discussion

### 1. Discussion of the English Performance Assessment Practiced by Science Teachers in CLIL Classrooms

Based on this study findings, English performance assessments in CLIL classrooms were mainly conducted using written tests and teacher observations. The written tests were designed with visual aids to support students and were used to evaluate writing, reading, and vocabulary skills. On the other hand, teacher observation was carried out spontaneously during classroom interactions without standardized criteria, aiming to assess students' oral and listening skills through activities like questioning, probing, and group discussions.

The types of assessments applied aligned with the two main categories suggested by Pokrivčáková (2010): formative and summative assessments. Teachers used written tests as summative tools at the end of the semester, while daily classroom observations functioned as formative assessments. This dual method of evaluation is also consistent with Rohmah (2022), who noted that the types of assessment commonly used by teachers are summative and formative, which serve different functions during and after the learning process. Similarly, Massler et al. (2011) emphasized that assessment in CLIL should include continuous observation and performance tasks to support both learning and teaching.

However, discrepancies were found in the assessment design, particularly regarding the written tests. The focus of scoring rubrics did not align with the intended learning outcomes in writing, emphasizing content over language accuracy. Furthermore, the written tests failed to cover



listening skills, contrasting with the view of Rohmah et al. (2019) who assert that CLIL assessments should address all language competencies, including listening and speaking, as part of comprehensive performance evaluation. As Short (1993) noted, the core challenge in assessment lies in separating language proficiency from content knowledge to prevent biased outcomes.

Observations conducted to assess speaking and listening were generally unstructured, yet they reflected techniques such as questioning and probing, aligning with Fitriati (2016) who highlighted the use of questioning strategies in classroom assessment. Despite their importance, observations were criticized for their lack of systematic implementation. Wewer (2014) pointed out that assessment practices in CLIL are often implicit and rely heavily on teachers' impressions, leading to concerns about reliability. Furthermore, Rohmah et al. (2019) revealed that without structured assessment instruments, teachers may overlook key student demonstrations, resulting in incomplete evaluation. This emphasizes the necessity for assessments that systematically integrate both content knowledge and language competencies. Supporting this, Rohmah (2019) found that such integration is not only feasible but also effective in enhancing students' English proficiency in CLIL settings, underscoring the need for well-designed assessment tools that address both dimensions.

To conclude, while the assessment methods showed some alignment with learning goals, several issues emerged. The assessment criteria for writing were inconsistent with expected outcomes, and listening was not evaluated in written tests. Scholars like McKay (2006) and Massler et al. (2011) emphasize that assessment should align with learning objectives and be supported by clear documentation to ensure validity. Additionally, Wewer (2014) underscored the necessity of systematic recording and standardization to produce reliable data, supporting the idea that proper documentation is critical to effective performance-based assessment.

## 2. Discussion of the Barriers Faced by the Science Teacher in Practicing English Performance Assessment in CLIL Classrooms

The findings of this study revealed several significant challenges faced by the science teacher and the co-teacher in implementing English performance assessments within a CLIL (Content and Language Integrated Learning) setting. These issues were primarily categorized into three major concerns: the underdevelopment of assessment criteria, the lack of appropriate instruments for language assessment, and the absence of effective strategies for assessing language skills. These challenges highlight the

difficulty the teachers faced in designing, applying, and evaluating English language performance alongside content instruction.

One of the central issues identified was the absence of clear, specific, and well-aligned criteria for assessing language skills. The teacher and co-teacher were unable to formulate assessment criteria that directly reflected the language learning objectives. As a result, the assessment tools used were vague and ineffective. This finding contradicts Lo (2014), who emphasized the importance of grounding assessment criteria in learning objectives to guide assessment instrument design. Although some attention was paid to aspects like pronunciation and vocabulary, the criteria failed to comprehensively assess all key language skills such as speaking, listening, reading, and writing. This runs counter to the principles outlined by Coyle et al. (2010) and Gablasova (2014), who argue that both content and language goals must be integrated in CLIL assessments.

This study further uncovered that the instruments used for assessment were not designed to evaluate language development. Instead, they focused almost entirely on measuring content achievement. This imbalance echoes the concerns raised by Short (1993), who highlighted the frequent issue in CLIL contexts where content dominates and language goals are marginalized. While Gass and Mackey (2007) and Coyle et al. (2010) have stressed the need to integrate both content and language objectives in assessment tools to support holistic learning, the teacher and co-teacher in this study struggled to reflect these dual aims in their practices.

Document analysis also revealed that the tools in use failed to monitor students' language progress. Rather than supporting holistic development, the assessments served mainly to test subject matter knowledge. According to Massler et al. (2011), CLIL assessment should encompass students' strategic use of language, content understanding, and even intercultural competence. However, in this case, important language features such as grammar, vocabulary usage, and genre awareness were not addressed in the assessment tools.

Another key finding was the teachers' limited use of assessment strategies that foster language development. There was a clear lack of methodological readiness, as both the teacher and co-teacher appeared unfamiliar with CLIL-specific assessment approaches. Massler (2011) proposed a five-step model for formative assessment in CLIL, involving varied tools like projects, visual aids, and language-support techniques, yet the teacher in this case relied mostly on conventional written tests, which are more suitable for content evaluation than for language assessment.

This reflects Rohmah et al. (2019), who emphasized that well-prepared CLIL implementation requires organized syllabi, varied materials, and a clear understanding of CLIL goals by the teachers. A contributing factor to this problem was the limited collaboration between the teacher and co-teacher. The responsibility for assessment planning and implementation rested mainly with the science teacher, while the English co-teacher played a minimal role. This lack of coordinated effort contrasts with the collaborative practices promoted by Heritage (2007), who emphasized the role of peer and teacher collaboration in formative assessment. Furthermore, Coyle et al. (2010) advocated for joint planning between content and language teachers to ensure that language demands are clearly addressed in instruction and assessment. In this study, the absence of such collaboration weakened the alignment between teaching goals and assessment practice.

## CONCLUSION

This study aimed to explore how English performance is assessed in CLIL (Content and Language Integrated Learning) classrooms and to identify the challenges faced by teachers in the assessment process. The findings show that the assessment of English performance in CLIL classrooms is conducted by science teachers using two primary methods: written tests and teacher observations. The written tests, typically administered at the end of the semester, are designed to evaluate students' reading, writing, and vocabulary skills through English-based test items that measure both content knowledge and language proficiency. These tests are usually prepared collaboratively in Teachers' Group Discussion forums to align with curriculum standards. In addition to the formal tests, teachers use ongoing informal observations during classroom activities such as questioning, group discussions, and presentations to assess students' speaking and listening skills. Although no formal observation instruments are used, teachers focus on how students use English in these contexts, reflecting the integrated nature of CLIL, where both content and language are assessed simultaneously.

However, the assessment process also presents several challenges. Teachers often lack clear and structured criteria for evaluating language skills, making it difficult to assess aspects such as pronunciation, grammar, vocabulary, and students' ability to express ideas in English. This is further complicated by the absence of specific assessment tools tailored for language proficiency, as most existing instruments primarily target content understanding. Additionally, limited use of effective assessment strategies, such as formative assessments, hinders the ability to monitor students' language development over time. These issues highlight the need for more structured assessment tools and teacher training to support the evaluation of English performance in CLIL classrooms. Enhanced collaboration between content

and language teachers is also essential to ensure that both content and language objectives are addressed effectively within the CLIL framework.

## REFERENCES

- Ball, D., & Lindsay, L. (2010). An investigation into CLIL-related sections of EFL course-books: Issues of CLIL inclusion in the publishing market. *International Journal of Bilingual Education and Bilingualism*, 17(3), 345–359.
- Bryman, A. (2006). Integrating quantitative and qualitative research: How is it done? *Qualitative Research*, 6(1), 97–113. <https://doi.org/10.1177/1468794106058877>
- Coyle, D., Hood, P., & Marsh, D. (2010). CLIL. *Cambridge University Press*. <https://doi.org/10.1017/9781009024549>
- Creese, A. (2010). Content-focused classrooms and learning English: How teachers collaborate. *Theory Into Practice*, 49(2), 99–105. <https://doi.org/10.1080/00405841003626494>
- Creswell, J. W. (2007). *Qualitative inquiry and research design: Choosing among five approaches* (2nd ed.). Sage Publications.
- Eurydice. (2006). Content and language integrated learning (CLIL) at school in Europe. Eurydice – The Information Network on Education in Europe.
- Fitriati, S. W. (2016). Safetalk practices in content and language integrated learning (CLIL) classes. *Language Circle: Journal of Language and Literature*, 11(1), 47–58. <https://doi.org/10.15294/lc.v11i1.7848>
- Gablasova, D. (2014). Learning and retaining specialized vocabulary from textbook reading: Comparison of learning outcomes through L1 and L2. *Modern Language Journal*, 98(4), 976–991. <https://doi.org/10.1111/modl.12150>
- Gass, S. M., & Mackey, A. (2007). *Input, interaction and output in second language acquisition*. In J. Williams & B. VanPatten (Eds.), *Theories in second language acquisition* (pp. 75–199).
- Heritage, M. (2007). Formative assessment: What do teachers need to know and do?. *Phi Delta Kappan*, 89(2), 140–145. <https://doi.org/10.1177/003172170708900210>
- Hüttner, J., & Rieder-Bünemann, J. (2010). Introduction: Content and foreign language integrated learning: A plurilingual perspective. In Y. Ruiz de Zarobe, J. M. Sierra, & F. Gallardo del Puerto (Eds.), *Content and foreign language integrated learning: Contributions to multilingualism in European contexts*. Peter Lang.
- Leech, N. L., & Onwuegbuzie, A. J. (2007). An array of qualitative data analysis tools: A call for data analysis triangulation. *School Psychology Quarterly*, 23(4), 557.

- Lo, Y. Y. (2014). Collaboration between L2 and content subject teachers in CBI: Contrasting beliefs and attitudes. *RELC Journal*, 45(2), 181-196. <https://doi.org/10.1177/0033688214535054>
- Massler, U., Ioannou-Georgiou, S., & Steiert, C. (2011). Effective CLIL teaching techniques: Guidelines for CLIL implementation in primary and pre-primary education.
- McKay, G. (2006). Task and performance-based assessment. In *Encyclopedia of language and education* (pp. 111-122).
- Miles, M. B., & Huberman, A. M. (1994). *An expanded sourcebook: Qualitative data analysis* (2nd ed.). Sage Publications.
- Pokrivčáková, S. (2010). *Modern teacher of English*. ASPA.
- Puspitasari, A., Anugerahwati, M., & Rachmajanti, S. (2016). Teachers' pedagogical and professional competences in CLIL-based primary schools in Indonesian context.
- Rachmajanti, S., Apriana, A., & Zen, E. (2015). Using English in science class as an attempt to maximizing input in second language learning. SEAMEO Regional TESOL Conference 2015. <http://www.vnseameo.org/TESOLConference2015/>
- Rohmah, I. I. T. (2019). The feasibility and effectiveness of integrating content knowledge and English competences for assessing English proficiency in CLIL. *ETERNAL (English Teaching Journal)*, 10(1),
- Rohmah, I. I. T. (2022). CLIL assessment: What every CLIL teachers should know. *Perkumpulan Rumah Cemerlang Indonesia*.
- Rohmah, I. I. T., Faridi, A., Saleh, M., & Fitriati, S. W. (2019). Content and language integrated learning (CLIL): Teachers' point of view. *European Union Digital Library*, 594-600. ISSN 2593-7650.
- Rohmah, I. I. T., Nurdianingsih, F., & Zainudin, M. (2021). The needs of standardized CLIL assessment in Indonesia. *Prosiding Nasional Pendidikan: LPPM IKIP PGRI Bojonegoro*, 2(1), 32-41.
- Rohmah, I. I. T., Saleh, M., Faridi, A., & Fitriati, S. W. (2019). Language assessment pattern for primary education in the content and language integrated learning (CLIL) classroom context. *Asian EFL Journal*, 21(2.2), 101-123.
- Ross, S. (2005). The impact of assessment method on foreign language performance growth. *Applied Linguistics*, 26(3), 317-342. <https://doi.org/10.1093/applin/ami011>.
- Short, D. J. (1993). Assessing integrated language and content instruction. *TESOL Quarterly*, 27(4),

Wewer, T. (2014). Assessment of young learners' English proficiency in bilingual content instruction CLIL.

Yin, R. K. (2003). *Case study research: Design and methods*. Sage Publications.