



VNU Journal of Foreign Studies

Journal homepage: <https://jfs.ulis.vnu.edu.vn/>



FROM TRAINING TO SUSTAINABILITY: A STUDY ON THE LONG-TERM IMPACT OF A CLIL PROGRAM ON TEACHER PRACTICE

Phan Thi Toan*

*Faculty of English Language and Culture, VNU University of Languages and International Studies,
No.2 Pham Van Dong, Cau Giay, Hanoi, Vietnam*

Received 28 April 2025

Revised 27 August 2025; Accepted 15 December 2025

Abstract: Building upon previous research investigating a professional development (PD) program designed to train primary school teachers in Content and Language Integrated Learning (CLIL) methodology, this follow-up study examines the long-term impact of the intervention approximately one year after completion. It aimed to determine the extent to which primary school teachers retain and implement CLIL principles after completing the PD program. The data were collected via multiple instruments, including a CLIL knowledge quiz, classroom observations, lesson plan analysis, and semi-structured interviews. The findings indicated strong short-term retention of the Content and Communication components, with challenges persisting in the implementation of Cognition and Culture. Contextual factors such as peer collaboration, student readiness, and time constraints played a critical role in shaping long-term adoption. The study contributes to understanding teacher learning as a situated, iterative process, and offers implications for designing sustainable CLIL-oriented PD.

Keywords: Content and Language Integrated Learning (CLIL), professional development, long-term impact, follow-up study

* Corresponding author.

Email address: toanphan.felte@gmail.com

<https://doi.org/10.63023/2525-2445/jfs.ulis.5505>

TỪ BỒI DƯỠNG ĐẾN BỀN VỮNG: KHẢO SÁT TÁC ĐỘNG CỦA CHƯƠNG TRÌNH CLIL ĐỐI VỚI VIỆC ÁP DỤNG DÀI HẠN CỦA GIÁO VIÊN

Phan Thị Toán

*Khoa Ngôn ngữ và Văn hóa Anh, Trường Đại học Ngoại ngữ, Đại học Quốc gia Hà Nội,
Số 2 Phạm Văn Đồng, Cầu Giấy, Hà Nội, Việt Nam*

Nhận bài ngày 28 tháng 4 năm 2025

Chỉnh sửa ngày 27 tháng 8 năm 2025; Chấp nhận đăng ngày 15 tháng 12 năm 2025

Tóm tắt: Tiếp nối nghiên cứu trước về một chương trình bồi dưỡng chuyên môn (PD) nhằm đào tạo giáo viên tiểu học theo phương pháp Tích hợp Nội dung và Ngôn ngữ (CLIL), nghiên cứu này khảo sát tác động dài hạn của can thiệp sau khoảng một năm kết thúc chương trình. Mục tiêu nghiên cứu nhằm xác định mức độ giáo viên tiểu học ghi nhớ và ứng dụng các nguyên lý của CLIL sau khi tham gia chương trình bồi dưỡng. Dữ liệu được thu thập thông qua nhiều công cụ, bao gồm: bài kiểm tra kiến thức CLIL, quan sát lớp học, phân tích giáo án và phỏng vấn bán cấu trúc. Kết quả cho thấy khả năng ghi nhớ ngắn hạn đối với hai thành phần: Nội dung (Content) và Giao tiếp (Communication) đạt mức cao, trong khi việc triển khai thành phần: Nhận thức (Cognition) và Văn hóa (Culture) vẫn còn gặp nhiều thách thức. Các yếu tố bối cảnh như: sự hỗ trợ đồng nghiệp, năng lực ngôn ngữ của học sinh và hạn chế về thời gian đóng vai trò quan trọng trong việc duy trì việc áp dụng CLIL lâu dài. Nghiên cứu đóng góp vào việc hiểu biết quá trình học hỏi của giáo viên như một tiến trình gắn với bối cảnh và mang tính tuần hoàn, đồng thời đề xuất các hàm ý cho việc thiết kế các chương trình bồi dưỡng CLIL bền vững.

Từ khóa: CLIL, phát triển chuyên môn, tác động dài hạn, nghiên cứu tiếp nối

1. Introduction

1.1. Statement of the Problem and Rationale

In Vietnam, there has been a long-term project to improve the quality of English teaching and learning throughout the country. The project, called Teaching and Learning Foreign Languages in the National Education System, period 2008–2020, was launched in 2008, and was set to last until 2020, then extended for five more years, until 2025 (MOET, 2016). It aimed to help any school leavers achieve CEFR level B1 in their chosen foreign language(s) by 2020.

Over several years, the Vietnam Ministry of Education and Training (MOET) has implemented various initiatives aimed at achieving national targets in teaching and learning. Since 2018, while maintaining improved educational outcomes as the ultimate goal, the focus has shifted to teacher professional development, raising teaching standards, leading many teachers nationwide to have opportunities to try new approaches and improve their pedagogical skills (Pham & Nguyen, 2020). MOET (2016) considers professional development crucial to the success of language teaching reform, thereby holding training programs every year for teachers to be trained and to practice new teaching methods, aiming at improving the learning experience for students.

One of these programs was conducted by the University of Languages and International Studies, Vietnam National University, Hanoi in 2023 and launched in several northern provinces

to help teachers in these areas to acquire knowledge of CLIL (Content and Language Integrated Learning). In this program, the theory of 4Cs (Content, Cognition, Communication and Culture) (Coyle et al., 2010) has been introduced, and during an eight-day training project (see Appendix), teacher trainees practiced applying the 4Cs into their real teaching. All training sessions were conducted in the summer over a short period of time. Their effectiveness was then measured through micro-teaching sessions, in which trainee teachers worked with their colleagues, who pretended to be their students to conduct lessons applying the target teaching approach. Although convenient and flexible, this assessment method potentially overlooks the pedagogical challenges and contextual factors influencing teacher performance in authentic classroom settings.

Teacher professional development (PD) has been a concern in many studies in recent years, and the ultimate objective of all teacher training programs has seemed to be for the teachers to effectively apply their knowledge in practice to support learning (Avalos, 2011). Evaluating the impact of PD initiatives is crucial for understanding their effectiveness in fostering sustained changes in teaching contexts. To evaluate the effectiveness of such programs, a considerable number of studies have been conducted (Darling-Hammond et al., 2017; King, 2013; Sims & Fletcher-Wood, 2020; Sims et al., 2023; Zeggelaar et al., 2020) over a range of contexts to see whether teachers could maximize the support to their students with what they have learnt. However, many such studies involve large participant numbers or assess impact immediately post-intervention, potentially limiting their ability to examine the nuanced, long-term effects of PD on individual teaching practices in specific contexts.

The current study, therefore, employs classroom observation and lesson plan analysis to investigate the integration of the CLIL approach into the participants' actual pedagogical practices, aiming to address the critical question of sustainability in teacher PD: to what extent do teachers retain and implement the principles learned in a short-term training program over the long term? It examines how contextual factors and a lack of ongoing support influence the teachers' ability to integrate CLIL into their daily practice. By exploring the gap between theoretical knowledge and practical implementation, this study contributes to the literature on effective teacher PD and offers insights for designing more sustainable CLIL training programs in the Vietnamese context.

1.2. Objectives of the Study

The study was expected to capture participants' perspectives and experiences regarding the changes in the teaching practice as the impact of the training program after one or several semesters. Moreover, any claims on the effectiveness of the PD program can be evidenced by the levels they retain the theory on the implementation of the approach.

The study seeks to answer the following questions:

- 1. To what extent do teachers retain and integrate CLIL theory into their teaching over time following the professional development intervention?*
- 2. What factors influence the long-term retention and implementation of CLIL principles among teachers who attended a PD program?*

There might be, apparently, other aspects that tend to be influenced by the project such as teacher efficacy or teacher autonomy. However, this study only focuses on teaching practices and teacher understanding of the theory. Building upon previous research that investigated a PD program designed to train primary school teachers in CLIL methodology, this article targets aims to present the findings of a follow-up study conducted approximately one year after the program's completion.

1.3. The Structure of the Article

This article is structured to provide a comprehensive overview of the follow-up study. It encompasses six main sections: Introduction, Theoretical Framework, Methodology, Findings, Discussion and Conclusion.

Introduction: provides the research background and context, and states the objectives of this follow-up study.

Theoretical Framework: presents a theoretical framework for effective PD along with a summary of the previous study, including the CLIL theory and the key findings.

Methodology: details the research design, participants, data collection tools, and the process of data analysis.

Findings: presents the results from the CLIL knowledge quiz, lesson plan analyses, classroom observations, and semi-structured interviews.

Discussion: includes the interpretation of the findings in the context of existing literature on effective PD and CLIL implementation, highlighting the persistent gaps between theory and practice.

Conclusion: summarizes the main findings of this study, acknowledges its limitations, and provides implications for future research and the design of more sustainable CLIL-oriented teacher training programs.

2. Theoretical Framework

2.1. Teacher Professional Development

Teacher PD is defined as teacher learning, which can occur in a wide range of formal and informal contexts, either planned or unplanned. Collinson et al. (2009) claim that nowadays, teachers who pursue continuous learning are necessary for social development, while Bakkenes et al. (2010) believe that school might not facilitate successful innovation because of the lack of teachers' learning. Moreover, research demonstrates that teacher learning impacts student learning (Buczynski & Hansen, 2010; Postholm, 2011a, 2011b), as reflected in higher exam scores, particularly when multiple teachers from the same school participate in PD programs.

Designed to facilitate learning, PD does not always result in meaningful professional growth. While positive effects on teachers and students have been observed, Bakkenes et al. (2010) indicate that teaching practices are the least influenced. They argue that a one-year period is insufficient for changes in knowledge and beliefs – typically the primary focus of PD research – to translate into changes in actual teaching practices. This view is shared by Ono and Ferreira (2010), who maintain that teachers require time to adjust their teaching and apply their PD learning effectively. Likely, Copur-Gencturk and Papakonstantinou (2016) maintain that PD programs often fail to produce the intended changes in teachers' instructional practices, which might be explained by the automatic activation of habitual behaviors triggered by environmental stimuli, as evidenced by a meta-analysis of causal studies across various contexts (Webb & Sheeran, 2006). Therefore, coaching where key features recognized for facilitating habit transformation are integrated could foster changes in teaching practice (Sims & Fletcher-Wood, 2020). On the other hand, Timperley et al. (2007) finds that standardized programs implemented without teacher involvement rarely yield lasting improvements in student outcomes, whereas those developed in collaboration with teachers are more effective and sustainable. Moreover, PD programs that incorporate instructional coaching – an ongoing cycle

of observation, feedback, and practice where teachers receive guidance from an experienced mentor – consistently demonstrate positive links to student achievement (Sims & Fletcher-Wood, 2020). Briefly, the structured cycle of repeated review and feedback within coaching models plays a crucial role in reinforcing and solidifying new cue–response relationships (Sims & Fletcher-Wood, 2020).

PD might fail because of several reasons. One of the reasons stated by Sims et al. (2023) is that teachers could not improve their understanding (insight) of teaching and learning through these programs. Furthermore, Sims et al. (2023) suggest that teachers may lack motivation to implement changes in their practice aligned with the content of PD programs. Cohen et al. (2020) also claim that simply knowing classroom management strategies seems to be insufficient to effect changes in teachers' practice, maybe because they are not provided with substantial skills to apply their acquired knowledge (Sims et al., 2023). The application might not even be long lasting without appropriate methods to embed such techniques into teachers' repertoire until they become the routine in their teaching (Sims et al., 2023).

In Vietnam, PD for teachers is widely acknowledged as a key aspect of language education, particularly with the nationwide implementation of the National Foreign Language Project 2020 (NFLP 2020). However, there is limited understanding of teachers' experiences with professional learning in their schools, the extent to which professional learning practices have evolved, and the impact of these changes on their development. Following a top-down approach and overseen by authorities, PD activities in Vietnam are typically conducted in designated locations with predetermined content (Nguyen, 2017; Nguyen & Mai, 2018). These activities take multiple forms, such as summer schools, qualification upgrading programs, demonstration lesson training, and in-school training (Le & Yeo, 2016; Nguyen, 2017; Saito, Khong & Tsukui, 2012). However, these models have been criticized for their inefficacy due to constraints related to time, funding, and human resources (Nguyen & Mai, 2018). Teachers are required to travel periodically to predetermined venues for PD workshops, which often fail to address their context-specific pedagogical challenges, as they do not consider teachers' individual needs or their school environments (Nguyen & Mai, 2018). Nguyen and Bui (2016) also imply that PD activities in Vietnam do not consider regional disparities or offer teachers adequate professional support and resources across different regions. Moreover, PD content is often characterized as theoretical, formal, and misaligned with teachers' practical needs (Le & Yeo, 2016; Nguyen, 2017). On top of that, teachers in Vietnam are not often willing to change their teaching habits (Le & Yeo, 2016) and do not usually receive sufficient support in their communities (Nguyen et al., 2019). Additionally, professional learning may be undervalued due to teachers' perception that certificates or awards are primarily needed to meet administrative requirements rather than being driven by their intrinsic motivation and learning needs (Le & Yeo, 2016). Despite significant financial investment in teacher PD through the NFLP 2020, it appears that teachers' learning needs were largely unmet. Teacher learning activities occurred largely in isolation, with minimal support provided (Nguyen et al., 2019). Even if the PD programs are effective, opportunities for participation were limited and not equally accessible to all teachers (Nguyen et al., 2019).

Regarding the PD of primary English language teachers, it has been insufficiently studied in the context of Vietnam (Nguyen & Pham, 2019). Their PD also tends to be affected by a number of factors, namely equal opportunities, relevance to their needs, opportunities for active learning and financial support (Nguyen & Pham, 2019). Besides, Nguyen and Pham (2019) claim that the teachers need more instructions with teaching pedagogy, particularly through real-time micro-teaching sessions, as well as additional demonstrations led by qualified master trainers and their colleagues.

Effective PD is organized learning designed to enhance teaching practices and lead to improved student learning outcomes (Darling-Hammond et al., 2017). To ensure PD is effective, various scholars have examined key factors that contribute to meaningful teacher learning. Wenger (2000) emphasizes the role of collaborative learning communities, indicating that significant and lasting improvements in teaching occur when educators collaborate within a professional community dedicated to fostering such changes. Opfer and Pedder (2011) emphasize that educational institutions must foster professional learning and engage teachers in shared knowledge and values, although this potentially necessitates additional support.

Building on comprehensive reviews of PD effectiveness, Desimone (2009) identifies five characteristics of effective PD. First, it must focus on content, including both subject matter and pedagogical knowledge. Second, learning should be active, incorporating observations, interactive feedback, and discussion. Third, PD content should align with teachers' prior knowledge and beliefs. Fourth, development activities should last at least one semester, with participant interaction exceeding 20 hours. However, Zwart et al. (2009) argue that even a one-year program may be too short for systematic changes, while Yoon et al. (2007) suggest that PD should range from 30 to 100 hours. In contrast, Basma and Savage (2017) and Kraft et al. (2018) claim that there is no link between the length of PD programs and their impact exerted on pupil achievement. Instead, it might be the repeated practices that are significant. Finally, teachers should learn collaboratively within a community and receive support from teacher leaders.

As identified by Darling-Hammond et al. (2017, p.4), seven characteristics of effective PD are that PD program:

1. Is content (that teachers teach) focused
2. Incorporates active learning utilizing adult learning theory: The structure of PD experiences should consider both the process of teacher learning and the content being learned.
3. Supports collaboration, typically in job-embedded contexts
4. Uses models and modeling of effective practice: video or written cases of teaching, demonstration lessons, unit or lesson plans, observations of peers, and curriculum materials including sample assessments and student work samples
5. Provides coaching and expert support
6. Offers opportunities for feedback and reflection
7. Is of sustained duration

2.2. Summary of the Previous Study

2.2.1. The 4Cs Framework of CLIL

The initial study was based on an 8-day PD program (including both theory and practice sessions) organized by the University of Languages and International Studies in 2023. The program aimed to introduce primary school teachers in Lạng Sơn province to the 4Cs framework of CLIL, as conceptualized by Coyle (2010). This framework served as the theoretical lens for analyzing teachers' lesson plans and classroom practices. The framework consists of four core components:

Content: The subject matter knowledge being taught (e.g., Science, Math).

Communication: The language used for learning and interaction in the classroom.

Cognition: The development of higher-order thinking skills.

Culture: The promotion of intercultural understanding and global citizenship.

2.2.2. Key Findings

The previous study (Phan, 2023) employed micro-teaching sessions and semi-structured interviews to analyze the extent to which teachers had absorbed and were able to apply these principles immediately following the course. It revealed a significant gap between the theoretical knowledge teachers acquired and their ability to translate it into practical application, particularly concerning the more complex CLIL components.

Firstly, teachers demonstrated a good grasp of the theoretical principles of Content and Communication. In their lesson plans, they clearly integrated subject matter knowledge and language objectives. During the micro-teaching sessions, teachers effectively used visual aids and gestures to convey content, and their lessons were engaging, with active student participation in communicative activities. This suggests that teachers were successful in mastering the foundational aspects of CLIL.

However, Cognition and Culture presented the most significant challenges for the teachers. Although the training program introduced these concepts, their practical application remained elusive. Teachers admitted to a deep confusion in designing activities that would foster higher-order thinking skills (e.g., analysis, evaluation) and found it difficult to naturally integrate the Culture component into their lessons. In the micro-teaching sessions, these elements were often either overlooked or addressed in a superficial, disconnected manner.

The interviews revealed that this gap was primarily due to a lack of practical experience and a perceived lack of confidence with a new pedagogical approach. Teachers acknowledged that the short duration of the training was insufficient for them to internalize the complex principles of CLIL and transform theoretical knowledge into fluent, practical teaching skills.

In conclusion, the initial study found that while the PD program successfully shifted teachers' mindsets and provided a positive conceptual foundation for CLIL, achieving sustainable application of all four CLIL components remained a significant challenge. These findings provide a crucial baseline for the present follow-up study, which investigates whether these challenges persist over the long term and how contextual factors have influenced teachers' implementation.

3. Methodology

3.1. Research Design

This study adopted a qualitative multiple-case study design (Yin, 2018) to explore how primary school teachers retained and implemented CLIL principles following a PD program. A case study approach was selected to allow an in-depth examination of individual teachers' understanding, pedagogical decision-making, and classroom practices over time.

The data collection process was structured in two main stages: a short-term assessment immediately after the PD program (the basis of the previous study) and a follow-up assessment approximately one year later. In the follow-up stage, by using a multi-method data collection strategy – including a CLIL knowledge quiz, lesson plan analysis, classroom observations, and semi-structured interviews – the study aimed to capture both teachers' conceptual retention and their actual teaching practices. Data from the CLIL knowledge quiz assessed teachers' theoretical retention. Classroom observations and lesson plan analysis provided insights into their practical application of CLIL principles. Finally, semi-structured interviews offered a deeper understanding of the reasons behind the level of retention and observed practices, linking teachers' PD experiences to their current implementation. This multi-faceted approach allowed the researcher to compare what teachers knew (quiz), what they planned (lesson plans), what

they did (observations), and what they said (interviews), thereby providing a robust foundation for the discussion of factors influencing long-term implementation.

3.2. Context and Participants

This second period of the study still involves the five trainee teachers who participated in the previous period. These teachers are of different ages and are assigned to teach in different sites, in the rural areas or in town. They also have different teaching experience and enjoy different employment status.

Figure 1

Profiles of the Five Participants (Phan, 2023)

Teacher	Age	Gender	Years of service	Employment status	Geographical location of school	Course book used in class
T1	32	Male	11	Permanent	rural	Global success
T2	35	Female	12	Permanent	rural	Global success
T3	33	Female	12	Permanent	rural	I-learn Smart Start
T4	32	Female	11	Permanent	rural	Global Success
T5	28	Female	7	Permanent	town	Global Success

According to the participants, their students have limited language skills, being able to produce only individual words in Grade 3, basic sentences in Grade 4, and simple conversations (with guidance) in Grade 5 (Phan, 2023). Therefore, the students in the micro-teaching sessions were of higher English proficiency than their real students. This, combined with the lack of further guidance or feedback after the training, was reported by the teachers as a factor that caused difficulties in applying the new practices in their authentic teaching contexts (Phan, 2023).

After the PD program that the participants participated in 2023, there were no further instructions regarding CLIL approach provided for any of them. Each of the teachers came back to their usual work, with the same roles and expectations. They no longer worked with the experts or together for further feedback and reflection. However, all of them had to take part in one more training session in the summer of 2024, in which they were instructed to handle the themes and topics in their textbook (Global Success) for the fifth grade. These later sessions, while not focused on CLIL, were based on traditional teaching standards and often involved assessments by colleagues who had not participated in the 2023 CLIL program. The assessment criteria and the content of this subsequent training, therefore, did not reinforce the CLIL principles and created a potential conflict between the new pedagogical approach and the existing professional expectations.

3.3. Data Collection and Analysis

3.3.1. Data Collection

Four instruments were used for data collection: (1) classroom observation, (2) lesson plan analysis, (3) a short CLIL knowledge quiz, and (4) semi-structured interviews.

First, classroom observation was conducted to explore how teachers translate CLIL theory into actual teaching practice. Each teacher was observed during at least one lesson, using the

observation protocol adapted from CLIL-specific frameworks that were used in the previous period of the study (Phan, 2023). This tool focuses on dimensions such as the integration of content and language objectives, cognitive challenge, use of scaffolding, and learner interaction. Each 35-minute lesson was video-recorded either by the researcher or, when direct site visits were not feasible, by an assistant cameraman, ensuring minimal disruption to the teaching process. Observational data provide valuable insight into the practical application of CLIL strategies and could be triangulated with lesson plans and interview data for deeper interpretation.

In addition to observing teaching practice, lesson plan analysis was employed to investigate how CLIL principles were reflected in teachers' instructional design. Participants were asked to submit the lesson plans for the observed lessons. These plans were analyzed using a content analysis approach, with attention paid to the articulation of dual objectives (language and content), task design, and evidence of CLIL-informed pedagogy. The analysis helped reveal whether CLIL theory continues to inform teachers' planning decisions and curriculum implementation.

Third, a short multiple-choice quiz consisting of twelve items was administered at the beginning of each interview session. The purpose of this quiz was to assess the extent to which teachers recalled and understood key CLIL concepts introduced during the PD program. The quiz covers fundamental principles of CLIL such as the 4Cs framework, scaffolding strategies, language integration, and learner engagement techniques. Teachers completed the quiz individually in 5–10 minutes. The responses were scored and coded based on accuracy to identify patterns of conceptual retention across participants. The CLIL knowledge quiz was used as an initial quantitative tool to assess teachers' theoretical retention of key CLIL concepts. While it provided an overview of what knowledge was retained, it did not offer insights into why certain concepts were forgotten or how contextual factors influenced their application. These were later achieved through the semi-structured interviews as a follow-up qualitative tool.

Semi-structured interviews, finally, were conducted to understand teachers' perspectives, experiences, and the contextual factors influencing their long-term engagement with CLIL. Interview questions were designed to explore current perception of CLIL, challenges encountered in applying CLIL principles, and types of support or constraints present in their teaching contexts. The interviews also allowed teachers to elaborate on their quiz responses and provide deeper context, such as explaining why they struggled with specific concepts like Cognition and Culture, as well as explain their lesson planning or classroom strategies in more depth. All interviews were audio-recorded, transcribed, and thematically analyzed using both inductive and deductive coding.

This approach of triangulating the data from the quiz, lesson observation and plans, and interviews provided a more comprehensive and nuanced picture, enabling the researcher to interpret the findings and link the theoretical knowledge to the complexities of real-world implementation.

Data collection proceeded in three stages aligned to the research questions and designed for triangulation:

Stage 1 – Planning and enactment (RQ1): Teachers' lesson plans were analyzed and non-participant classroom observations were conducted to examine the extent to which CLIL was integrated from planning to real-time instruction.

Stage 2 – Knowledge retention (RQ1): A 12-item multiple-choice quiz captured teachers' retention and understanding of key CLIL principles (4Cs).

Stage 3 – Sense-making and PD linkages (RQ1–RQ2): Semi-structured interviews probed how teachers interpreted CLIL and how specific PD experiences (e.g., modeling, collaborative tasks, coaching, feedback, duration) supported or constrained implementation.

3.3.2. Data Analysis Framework

This study draws on Coyle et al.'s (2010) 4Cs framework of CLIL as the central analytical lens across all data sources: classroom observation transcripts, lesson plans, CLIL knowledge quiz, and semi-structured interviews. Using a consistent coding scheme grounded in these four dimensions enables a systematic analysis of both theoretical retention and practical implementation of CLIL among primary teachers following the PD intervention.

3.3.2.1. Coding Scheme

The coding scheme was developed to bridge the gap between abstract CLIL theory and observable classroom practices. The focus was on specific, well-documented activities that serve as tangible indicators of CLIL implementation. This approach allowed for a systematic and replicable analysis across all participants.

Content (CT): The study analyzed how subject-specific knowledge and skills were conveyed. For example, in the classroom observation transcripts, the researcher coded for the “use of disciplinary content in English (e.g., verbs of ability, body parts)”.

Cognition (CG): To assess the promotion of cognitive processing, the researcher looked for evidence of higher-order thinking skills, such as the “use of open-ended questions, reasoning tasks, and student-created responses”, which are central to the Cognition dimension.

Communication (CM): The researcher coded for how the target language was used for interaction, scaffolding, and output. Key indicators included “pair/group interaction in English, modeled language, and repetition”.

Culture (CU): To evaluate the integration of cultural awareness, the researcher looked for evidence of “personalization, cross-cultural discussion, and references to values or global themes”.

3.3.2.2. Data Analysis Process

The 4Cs coding scheme was applied consistently across all data sources to enable triangulation and validate our findings:

Classroom Observation Transcripts: Transcripts were coded line-by-line using the scheme above.

Lesson Plan Analysis: The researcher used a rubric based on the same 4Cs framework to assess each lesson plan for the presence of each CLIL element.

CLIL Knowledge Quiz: Quiz responses were analyzed using a High/Medium/Low scale for each 4Cs dimension, reflecting teachers’ conceptual understanding and interpretation.

Semi-Structured Interviews: Interview transcripts were coded thematically using the 4Cs as overarching categories to understand teachers’ perspectives and contextual factors.

This layer helped interpret variation in classroom practice and revealed how teachers internalized or adapted CLIL theory over time, thereby strengthening the trustworthiness of the findings and providing a comprehensive view of CLIL implementation and its sustainability.

4. Findings

4.1. Retention of CLIL Theory (RQ1)

The CLIL knowledge quiz served as the primary quantitative measure of theoretical retention. Teachers completed a 10-item multiple-choice test with two additional scenario-based questions. The results revealed very high levels of retention in all the areas, especially the ones of Content and Communication. Teachers demonstrated strong understanding in areas related to content-language integration, vocabulary development, the purpose of scaffolding, and the role of visuals, answering associated quiz items correctly. Notably, questions assessing the theoretical foundations of CLIL, such as the 4Cs and the integration of cultural understanding, also yielded perfect results. Culture-related questions and those that addressed cognition at higher levels (e.g., applying and evaluating) were the least (still very highly) accurately answered.

Interview data triangulated the quiz findings. They indicated a divergent pattern of CLIL retention among teachers. Most respondents were able to recall general principles of the 4Cs framework but tended to focus on the more tangible elements such as Content and Communication. For example, in response to Q1 and Q2, teacher T2 defined the easiest element to apply as *“vocabulary teaching and structured sentence use,”* reflecting retention of Content, while teacher T3 highlighted *“getting students to talk with support,”* pointing to Communication. Some teachers expressed uncertainty when asked to define or describe Cognition and Culture, indicating partial conceptual retention. For instance, Teacher T1 depicted CLIL as *“mostly about using English with content vocabulary,”* without mentioning the broader pedagogical goals such as fostering critical thinking or intercultural competence. A notable comment by teacher T2 in Q2 described *“tasks requiring higher-order thinking”* as challenging due to both language and cognitive demands. In contrast, Teachers T4 and T5 demonstrated a clearer grasp of CLIL principles, using terms like *“language objectives,” “support strategies,”* and *“student interaction.”*

Furthermore, when asked about the quiz itself, several teachers noted that while they remembered the theoretical terminology from the PD program, they had forgotten practical techniques. This divergence between conceptual and procedural memory suggests that although surface-level knowledge was retained, deeper theoretical understanding had eroded over time for some participants.

4.2. Integration of CLIL in Teaching Practice (RQ1)

4.2.1. Lesson Plan Analysis

The lesson plans submitted by the participating teachers were reviewed and evaluated against the 4Cs framework. These plans, while diverse in their topics and specific content, were found to share common objectives, reflecting the curriculum requirements for grades 3, 4 and 5. The topics analyzed included *“In the City,” “Food and Drinks,” “Toys”* and *“Health.”*

The overarching objectives across these plans focused on integrating language skills with practical, real-world knowledge. For instance, the *“In the City”* lesson aimed to teach students how to ask for and give directions, while the *“Health”* lesson focused on reviewing vocabulary to discuss health problems and offer advice. These plans served as a crucial tool for assessing how teachers intended to apply the CLIL principles they had learned, allowing the researcher to compare their levels of theory retention with their plans to implement it in real lessons.

Each plan was assessed using a rubric approach to identify the explicit or implicit

presence of each CLIL element. This method provided a systematic way to evaluate the level of CLIL integration across all participants.

Figure 2

Summary of Lesson Plan Analysis

Teacher	Content (CT)	Cognition (CG)	Communication (CM)	Culture (CU)
T1	High – Lessons include functional language on giving directions and road signs, tightly linked to real-life content	Medium – Students make predictions, match and complete tasks, but critical thinking remains basic	High – Pair work, structured dialogues, and interactive games encourage language use	Low – No mention of cultural or intercultural content
T2	High – Strong integration of functional phrases, signs, and directions in urban contexts	Medium – Students guess and complete gaps, but tasks remain within familiar routines	High – Many interactive activities including singing, role-plays, and peer questioning	Low – Culture not addressed explicitly
T3	High – Consistent vocabulary and sentence patterns linked to health, advice, and food	Medium – Students give examples, write and review, but tasks focus on application rather than analysis	Medium – Teacher-led questioning and individual/pair work; limited group-based communication	Low – No intercultural links made, though some topics are personally relevant
T4	High – Food-related vocabulary embedded in chant, listening, and role-play	Medium – Students describe comics, complete tasks, but no explicit reflection or reasoning	Medium – Use of chanting and pair work, but few collaborative or problem-solving activities	Low – Cultural dimension absent; no reference to food practices across cultures
T5	Medium – Vocabulary on toys and structures like “They have...” are well-presented but not expanded into broader contexts	Medium – Simple recall and completion tasks dominate; no deeper processing	Medium – Some paired speaking activities, but interaction mostly repetition-based	Low – No cultural comparison; content remains local and familiar

According to the table, all teachers showed strong evidence of content integration, with structured vocabulary input aligned to thematic units such as health, food, directions, and toys. Teachers T1 and T3 provided lesson plans that connected vocabulary directly to classroom tasks (e.g., giving directions using maps, naming food items in health lessons). Sentence frames, chants, and contextual examples were embedded to support comprehension.

Communication was also strongly represented across most lesson plans. Teachers designed activities that encouraged speaking and listening, with structured prompts, dialogues, and songs. Teacher T4, for example, included a chant about healthy eating followed by a role-play between students discussing food preferences.

Cognition was addressed to a medium extent in all lesson plans. Tasks such as matching, classifying, gap-filling, and yes/no questioning were prevalent, while fewer examples of higher-order thinking (e.g., evaluation, creation, synthesis) were observed. Teacher T1 stood out for

including an activity in which students predicted the next steps in a story and justified their answers.

Culture, although relevant and possible to be integrated in all lessons, remained the least integrated element. Across the lesson plans, only teacher T2 made an implicit reference to cultural norms through a lesson on direction giving. Most lessons stayed within local or textbook-defined contexts without extending to intercultural comparison or reflection.

4.2.2. Classroom Observation

Figure 3

Coded Class Observation Transcripts

Teacher	Content (CT)	Cognition (CG)	Communication (CM)	Culture (CU)
T1	High – Focused on the vocabulary related to road signs and directions	Medium – Students have to understand and apply using the tools for different tasks	High – There are pair and group activities where students use English to perform tasks together.	Low – There are few cultural connections in the lesson
T2	High – Clear use of content vocabulary and structure tied to lesson goals (e.g., directions, urban places)	Medium – Students guess, match, and recall information; higher-order thinking limited	High – Frequent pair/group practice with English used for real purposes	Low – Cultural content not explicitly integrated
T3	High – Systematic teaching of food vocabulary with corresponding visuals and characters	Medium – Students identify preferences and match audio content, but tasks are mostly comprehension-based	High – Use of songs, listening and drawing lines, personal responses in English	Low – Cultural connections not present beyond food items
T4	High – Lessons centered around food and vocabulary review with reading and writing practice	High – Involves listening discrimination, gap-filling, reasoning, and vocabulary application	High – Students speak in full sentences, read aloud, participate in games and discussion	Medium – Includes real-life context (e.g., food in celebrations in different countries), but cultural depth is limited
T5	Medium – Focused on vocabulary recognition (e.g., toys, colors) without extended content development	Medium – Uses matching and recall-based activities with limited problem-solving	High – Emphasis on chants, dialogues, and pair work	Low – No significant cultural reference or comparison

The observations reinforced the patterns found in lesson plans. As revealed, content delivery was strong, with teachers consistently embedding English vocabulary into content topics. For example, teacher T1 introduced names of road signs in English, while teacher T3 taught types of food using both visuals and oral drills. Teachers provided clear explanations, used gestures and repetition effectively, and structured lessons around target vocabulary.

Communication was observed through teacher-student interaction and, in several cases,

peer interaction. Teachers T2 and T4 facilitated group games and pair-work dialogues that allowed students to practice language in authentic contexts. However, in teachers T3 and T5's classes, interaction was predominantly teacher-led, and student responses were often limited to repetition or one-word answers. Moreover, both teachers T1 and T3 showed limited interactions with their students; they rarely left the standing stage to support students in individual or pair tasks.

Cognition appeared in comprehension checks, sequencing activities, and question-answer routines. While lower-level thinking (recall, understanding) dominated, some evidence of problem-solving and reasoning was present, particularly in teacher T4's use of open-ended questions and teacher T2's use of table analysis.

Culture was least observable. Apart from a brief discussion in teacher T4's class on different food in different countries, no significant intercultural elements were present. Although teacher T2 asked the students to recall their local traffic signs, the connections were not developed further.

The analysis of both lesson plans and classroom observations revealed patterns of alignment and divergence in the application of CLIL principles across the 4Cs framework. While teachers might plan with some CLIL-aligned strategies, the extent to which these were enacted in practice varied by component and teacher.

Content (CT) - Alignment: High

Across all teachers, the integration of content vocabulary and thematic focus was consistent between lesson plans and observed instruction. In lesson plans, content was often introduced through structured vocabulary sets and embedded within contextual tasks. This was mirrored in the classrooms where teachers used visuals, chants, and realia (e.g., maps, food pictures) to reinforce subject-specific vocabulary. This shows that teachers retained and successfully implemented content-language integration from the PD program.

Communication (CM) - Alignment: Moderate

Lesson plans typically included opportunities for communicative practice – dialogues, songs, sentence frames – and these were often realized in practice. However, there were discrepancies in the depth of student interaction. While teachers designed tasks involving pair or group work, classroom observations showed that these were sometimes replaced by teacher-fronted questioning or choral repetition. For example, Teacher T3's plan indicated pair interaction for discussing food preferences, but in class, students mostly repeated answers individually or responded to the teacher. This suggests that while the principle of using English for communication was retained, its execution lacked sufficient scaffolding or collaborative structure in some classrooms.

Cognition (CG) - Alignment: Low to Moderate

Cognitive tasks in lesson plans were mostly low-order (e.g., matching, filling blanks). In classroom observations, the same trend persisted – students engaged in comprehension and recall-based tasks, with only occasional examples of reasoning or justification. For instance, Teacher T4 included a prediction task in the lesson plan (*"What will happen next in the comic?"*) and asked students to explain their choices during the lesson – indicating alignment and application of cognitive scaffolding. This suggests that while cognition is understood, its application remains basic in both planning and instruction.

Culture (CU) - Alignment: Low

Both lesson plans and observations consistently showed minimal or no integration of

cultural content. Teachers rarely planned for or implemented activities that encouraged intercultural awareness or reflection. Teacher T5 mentioned diversity (e.g., “*different genders might like different toys*”) during the lesson, but this was not part of the lesson plan and was not extended as a learning goal. This indicates that the Culture component of CLIL remains the least internalized, highlighting a persistent gap between theory and practice.

This comparative analysis reveals that while Content and Communication were well planned and frequently implemented, Cognition and Culture were more vulnerable to reduction or omission in real-time teaching.

4.3. Influencing Factors on Retention and Implementation (RQ2)

The study revealed that factors external to the teachers, such as student proficiency, lack of support, and time constraints, significantly impacted their ability to implement CLIL principles. The interview data directly linked these challenges to their PD experiences and subsequent classroom practices.

For instance, teachers highlighted student proficiency as a major barrier. As Teacher T3 noted, “*It is a challenge to conduct high-order thinking activities with students whose English is not good enough.*” This perceived lack of student readiness led to a simplification of lessons, often abandoning more complex CLIL activities in favor of more basic language practice.

Similarly, a lack of institutional support was a critical factor. Teacher T1 stated, “*I have to do all by myself without any support from the school leader or my colleagues to deal with the challenges.*” This feeling of isolation contrasted with the collaborative environment of the PD program and hindered long-term adoption. The absence of a continuous support system meant teachers were left to navigate implementation challenges on their own. The only exception among the participants was teacher T4, who collaborated with peers and received administrative encouragement, and also who demonstrated higher levels of CLIL retention and integration.

Finally, curriculum pressure and time constraints were frequently cited. Teacher T1 commented, “*There is no time to develop special materials. I just follow the textbooks.*”, while teachers T2 and T3 admitted that they “*only focus on helping students to do well in the tests*”. This pressure to cover the curriculum for standardized tests meant that teachers prioritized content coverage over developing and implementing creative CLIL activities, which they had learned in the PD program. These findings show a clear discrepancy between what teachers were trained to do and what they were able to do in their actual teaching contexts.

Besides contextual constraints, emotional barriers also directly impact CLIL implementation over time. Through the interviews, it was indicated that teachers who viewed CLIL as aligned with their teaching values (e.g., Teachers T2, T4) incorporated it more fully. Others regarded it as an external requirement, so implemented it only superficially (Teachers T3, T5).

5. Discussion

This study’s long-term findings shed light on the complex relationship between a short-term professional development (PD) program and its lasting impact on teacher practice. This section will interpret the findings on teachers’ retention and implementation of CLIL principles, analyze the PD program itself in the context of effective PD literature, and conclude with the broader implications for designing sustainable teacher training in Vietnam.

The Persistent Gap between Knowledge and Practice

Our findings from both the initial and follow-up studies reveal a persistent and

significant gap between teachers' theoretical understanding and their practical application of CLIL. While the teachers demonstrated a robust grasp of the foundational Content and Communication components, challenges with Cognition and Culture remained. This discrepancy is evidenced by teachers' self-reported confusion and difficulties in applying these more abstract principles, a finding that is consistent with previous research on PD effectiveness. For instance, Teacher T3 stated, *"The culture element was mentioned and explained, but I don't know how to apply it in my class,"* suggesting that theoretical knowledge alone, without practical reinforcement and contextual support, is insufficient for successful implementation. This finding aligns with the literature on knowledge transfer, which suggests that complex skills require sustained practice and guided application to become fully integrated into one's professional repertoire (Webb & Sheeran, 2006).

The PD Program: A Necessary but Insufficient Step

When compared with established theories of effective PD, this CLIL program satisfied several key features. It ensured the content was directly relevant to what teachers teach at school and incorporated active learning with opportunities for collaboration and feedback, aligning with recommendations from Desimone (2009) and Yoon et al. (2007). However, these positive features were ultimately insufficient to guarantee long-term success.

The PD program's most critical limitation was the lack of sustained support. As a short, 8-day intervention, it did not provide the long-term, job-embedded support that is crucial for building new habits and skills. This directly addresses the teachers' voiced needs for *"more ready-made materials"* (teachers T1, T3, T5), *ongoing coaching* (teachers T2, T3, T4), and *opportunities to co-plan lessons* (teachers T1, T3), which are considered essential components of effective PD (Darling-Hammond et al., 2017; Sims & Fletcher-Wood, 2020). The absence of a post-training professional support infrastructure meant teachers were left to navigate contextual challenges on their own, often reverting to familiar practices. This highlights that a one-off PD event, no matter how engaging, is often inadequate for promoting deep, lasting pedagogical change (Timperley et al., 2007).

Furthermore, the findings suggest that the PD program failed to fully account for the diverse teaching contexts of the participants. While all teachers came from the same province, their teaching environments varied significantly. The lack of time for self-reflection and personalized support for these differing contexts signals the limited long-term success of a "one-size-fits-all" training approach, a point also made by Zeggelaar et al. (2020).

Mindset Shift

Despite these challenges, the study also uncovered a positive shift in teacher mindset and a renewed sense of professional agency. For example, Teacher T2 shared, *"Before that, I had had no idea of using English effectively in class instructions, but now I try to expose the students to the content with appropriate language."* This reflection suggests that the PD program successfully raised teachers' awareness of the value of CLIL and motivated them to make conscious changes in their classrooms, a critical first step in the behavior change process (Webb & Sheeran, 2006).

This finding adds a crucial layer of nuance to the understanding of the PD's impact. It suggests that while successful CLIL implementation requires a confluence of contextual, relational, and motivational supports, the initial PD program was a necessary catalyst for this mindset shift. This study's findings therefore position teacher learning as a situated and iterative process (Wenger, 2000), where initial training acts as a foundational step that must be followed by a sustained and collaborative professional learning journey.

6. Conclusion

6.1. Summary of the Findings

This study, as a follow-up to previous research, provides a unique and important perspective on the long-term impact of a CLIL professional development program for primary school teachers. Our findings indicate that while teachers demonstrated strong retention of the Content and Communication components of CLIL, the implementation of Cognition and Culture remained a significant challenge. This discrepancy was primarily influenced by contextual factors such as student proficiency, lack of institutional support, and time constraints.

These findings suggest that teachers need not only content knowledge but also confidence, collaboration, and creative autonomy. Deep theoretical understanding, while ideal, does not automatically translate to effective practice. What matters is the teacher's capacity to adapt strategies to learners' needs – a core principle of pedagogical content knowledge (Shulman, 1987). This reinforces the idea that implementation is not merely about memory, but about mindset, opportunity, and ongoing dialogue.

Moreover, short-term, one-off training programs may not be sufficient to bring about deep and lasting changes in teaching practice. Instead, effective PD should be seen as a continuous process that includes ongoing support, collaboration, and opportunities for reflection. Future PD programs in Vietnam should consider these findings and incorporate sustained support mechanisms to help teachers translate their newly acquired knowledge into sustainable classroom practice.

6.2. Limitations and Implications for Future Research

While this study offers valuable insights into the long-term impact of a CLIL PD program on primary school teachers in a mountainous province of Vietnam, several limitations should be considered when interpreting the findings.

Small Sample Size: The research focused on only five teachers. Although this case study approach allowed for in-depth analysis, the small sample size limits the generalizability of the findings to the broader population of primary school teachers within the province or other educational contexts in Vietnam and internationally.

Context Specificity: The findings may be influenced by the specific context of the mountainous province where the study was conducted, including factors such as students' English proficiency levels, available resources, and local educational policies. Therefore, caution should be exercised when applying the results to different regions.

Reliance on Qualitative Data: The study predominantly relied on qualitative data collection methods, including classroom observations, lesson plan analysis, and interviews. While these methods provide rich and detailed data, they can be subject to the researcher's subjective interpretation. The absence of more extensive quantitative measures (beyond the knowledge quiz) may limit a more objective assessment of the impact.

Observer Effect: The presence of the researcher or a cameraman during classroom observations might have inadvertently influenced the teachers' natural teaching behavior (the Hawthorne effect), potentially leading them to perform differently than they would under normal circumstances.

Lack of Control Group: The study did not include a control group of teachers who did not participate in the CLIL training program for comparison, similar to the limitation noted for the first part of the research. This makes it challenging to definitively attribute the observed

changes in teaching practices solely to the professional development intervention.

Timeframe: Although the study examined effects occurring six months to over a year post-training, this timeframe might still be insufficient to fully assess the sustainability and deep embedding of changes in teachers' habitual practices.

Acknowledging these limitations does not diminish the study's value but helps to situate the findings appropriately and suggests avenues for future research designed to address these constraints.

References

- Avalos, B. (2011). Teacher professional development in Teaching and Teacher Education over ten years. *Teaching and Teacher Education*, 27(1), 10–20. <https://doi.org/10.1016/j.tate.2010.08.007>
- Bakkenes, I., Vermunt, J. D., & Wubbels, T. (2010). Teacher learning in the context of educational innovation: Learning activities and learning outcomes of experienced teachers. *Learning and Instruction*, 20(6), 533–548. <https://doi.org/10.1016/j.learninstruc.2009.09.001>
- Basma, B., & Savage, R. (2017). Teacher professional development and student literacy growth: A systematic review and meta-analysis. *Educational Psychology Review*, 30(2), 457–481. <https://doi.org/10.1007/s10648-017-9416-4>
- Buczynski, S., & Hansen, C. B. (2010). Impact of professional development on teacher practice: Uncovering connections. *Teaching and Teacher Education*, 26(3), 599–607. <https://doi.org/10.1016/j.tate.2009.09.006>
- Cohen, J., Wong, V., Krishnamachari, A., & Berlin, R. (2020). Teacher Coaching in a Simulated Environment. *Educational Evaluation and Policy Analysis*, 42(2), 208–231. <https://doi.org/10.3102/0162373720906217>
- Collinson, V., Kozina, E., Kate Lin, Y., Ling, L., Matheson, I., Newcombe, L., & Zogla, I. (2009). Professional development for teachers: a world of change. *European Journal of Teacher Education*, 32(1), 3–19. <https://doi.org/10.1080/02619760802553022>
- Copur-Gencturk, Y., & Papakostantinou, A. (2016). Sustainable changes in teacher practices: A longitudinal analysis of the classroom practices of high school mathematics teachers. *Journal of Mathematics Teacher Education*, 19(6), 575–594. <https://doi.org/10.1007/s10857-015-9310-2>
- Coyle, D., Hood, P., & Marsh, D. (2010). *CLIL: Content and language integrated learning*. Cambridge University Press.
- Darling-Hammond, L., Hyler, M. E., & Gardner, M. (2017). *Effective teacher professional development*. Learning Policy Institute. <https://learningpolicyinstitute.org/product/effective-teacher-professional-development-report>
- Desimone, L. M. (2009). Improving impact studies of teachers' professional development: Toward better conceptualizations and measures. *Educational Researcher*, 38(3), 181–199. <https://doi.org/10.3102/0013189X08331140>
- King, F. (2013). Evaluating the impact of teacher professional development: An evidence-based framework. *Professional Development in Education*, 40(1), 89–111. <https://doi.org/10.1080/19415257.2013.823099>
- Kraft, M. A., Blazar, D., & Hogan, D. (2018). The effect of teacher coaching on instruction and achievement: A meta-analysis of the causal evidence. *Review of Educational Research*, 88(4), 547–588. <https://doi.org/10.3102/0034654318759268>
- Le, P. H. H., & Yeo, M. (2016). Evaluating in-service training of primary English teachers: A case study in Central Vietnam. *The Asian EFL Journal Quarterly*, 18(1), 163–191.
- Ministry of Education and Training (MOET). (2016, March 17). *Official Letter No. 45/CV-DANN on key tasks in 2016 of the National Foreign Languages Project 2020*. Thu Vien Phap Luat. <https://thuvienphapluat.vn/cong-van/Giao-duc/Cong-van-45-CV-DANN-2016-nhiem-vu-trong-tam-cua-De-an-Ngoai-ngu-Quoc-gia-2020-307261.aspx>
- National Foreign Language Project 2020 Management Board (NFLP 2020). (2016). *Teaching and learning foreign languages in the national education system, period 2008–2020: Results of the 2011–2015 period and implementation plan for the 2016–2020 period* [Paper presentation]. The NFLP 2020 Conference on Reviewing Outcomes 2011–2015 and Planning Strategies 2016–2020, Hanoi, Vietnam. <https://www.ctu.edu.vn/tin-tuc-su-kien/hoi-nghi-trien-khai-de-an-day-va-hoc-ngoi-ngu-trong-he-thong-giao-duc-quoc-dan-giai-doan-2008-2020.html>

- Nguyen, H. T. M. (2017). *Model of mentoring in language teacher education*. Springer. [10.1007/978-3-319-44151-1](https://doi.org/10.1007/978-3-319-44151-1)
- Nguyen, H. T. M., & Bui, T. (2016). Teachers' agency and the enactment of educational reform in Vietnam. *Current Issues in Language Planning*, 17(1), 88–105. <https://doi.org/10.1080/14664208.2016.1125664>
- Nguyen, H. T. M., Phan, H. L. T., & Le, N. D. (2019). Teachers' professional learning in the context of language education reforms. In H. T. M. Nguyen, H. L. T. Phan & N. D. Le (Eds.), *Teacher professional learning in the context of education reforms* (pp. 85-104). Routledge. [10.4324/9780429457371-6](https://doi.org/10.4324/9780429457371-6)
- Pham, Q. H. & Nguyen, N. D. (2020). Restructuring the teacher education system in Vietnam. *International Journal of Learning, Teaching and Educational Research*, 19(4), 29-43. [10.26803/ijlter.19.4.3](https://doi.org/10.26803/ijlter.19.4.3)
- Nguyen, T. T. T., & Pham, T. T. T. (2019). Primary English language teachers' engagement in professional development. *VNU Journal of Foreign Studies*, 35(1), 1-18. <https://doi.org/10.25073/2525-2445/vnufs.4342>
- Nguyen, V. T., & Mai, N. K. (2018). Professional development as part of English education initiatives in the ASEAN. In K. Hashimoto & V. T. Nguyen (Eds.), *Professional development of English language teachers in Asia: Lessons from Japan and Vietnam* (pp. 11–25). Routledge. <https://doi.org/10.4324/9781315413259>
- Ono, Y., & Ferreira, J. (2010). A case study of continuing teacher professional development through lesson study in South Africa. *South African Journal of Education*, 30(1), 59-74. <https://doi.org/10.15700/saje.v30n1a320>
- Opfer, V. D., & Pedder, D. G. (2011). The lost promise of teacher professional development in England. *European Journal of Teacher Education*, 34(1), 3–23. <http://dx.doi.org/10.1080/02619768.2010.534131>
- Phan, T. T. (2023). Applying CLIL in teaching at primary schools. In *Proceedings of the International Conference on New Trends in English Language Training at Multi-disciplinary and Interdisciplinary Universities* (pp. 339–364). Vietnam National University Press.
- Postholm, M. B. (2011a). Teachers' learning in a research and development work project. *Educational Action Research*, 19(2), 231–244. <https://doi.org/10.1080/09650792.2011.569237>
- Postholm, M. B. (2011b). A completed research and development work project in school: The teachers' learning and possibilities, premises and challenges for further development. *Teaching and Teacher Education*, 27(3), 560–568. <https://doi.org/10.1016/j.tate.2010.10.010>
- Saito, E., Khong, T. D. H., & Tsukui, A. (2012). Why is school reform sustained even after a project? A case study of Bac Giang Province, Vietnam. *Journal of Educational Change*, 13(2), 259–287. <https://doi.org/10.1007/s10833-011-9173-y>
- Shulman, L. S. (1987). Knowledge and teaching: Foundations of the new reforms. *Harvard Educational Review*, 57(1), 1-22. <https://doi.org/10.17763/haer.57.1.j463w79r56455411>
- Sims, S., & Fletcher-Wood, H. (2020). Identifying the characteristics of effective teacher professional development: A critical review. *School Effectiveness and School Improvement*, 32(1), 47-63. <https://doi.org/10.1080/09243453.2020.1772841>
- Sims, S., Fletcher-Wood, H., O'Mara-Eves, A., Cottingham, S., Stansfield, C., Goodrich, J., Van Herwegen, J., & Anders, J. (2023). Effective Teacher Professional Development: New Theory and a Meta-Analytic Test. *Review of Educational Research*, 95(2), 213-254. <https://doi.org/10.3102/00346543231217480>
- Timperley, H., Wilson, A., Barrar, H., & Fung, I. (2007). *Teacher professional learning and development: Best evidence synthesis iteration*. Wellington: Ministry of Education. <https://www.educationcounts.govt.nz/publications/series/2515/15341>
- Webb, T. L., & Sheeran, P. (2006). Does changing behavioral intentions engender behavior change? A meta-analysis of the experimental evidence. *Psychological Bulletin*, 132(2), 249–268. <https://doi.org/10.1037/0033-2909.132.2.249>
- Wenger, E. (2000). Communities of practice and social learning systems. *Organization*, 7(2), 225-246. <https://doi.org/10.1177/135050840072002>
- Yin, R. K. (2018). *Case study research and applications: Design and methods* (6th ed.). Sage Publications.
- Yoon, K. S., Duncan, T., Lee, S. W. Y., Scarloss, B., & Shapley, K. (2007). *Reviewing the evidence on how teacher professional development affects student achievement* (REL 2007-No. 033). U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Southwest. <http://ies.ed.gov/ncee/edlabs>

Zeggelaar, A., Vermeulen, M., & Jochems, W. (2020). Evaluating effective professional development. *Professional Development in Education*, 48(4), 665-682. <https://doi.org/10.1080/19415257.2020.1744686>

Zwart, R. C., Wubbels, T., Bergen, T., & Bolhuis, S. (2009). Which characteristics of a reciprocal peer coaching context affect teacher learning as perceived by teachers and their students? *Journal of Teacher Education*, 60(3), 243–257. <http://dx.doi.org/10.1177/0022487109336968>

APPENDIX

Training Course Outline for the 8-Day Summer CLIL PD Program

Day	Topics / Activities
1	<ul style="list-style-type: none"> • Introduction to the CLIL approach • The 4Cs model in CLIL • Writing content and language objectives for CLIL lessons • Adapting learning materials for CLIL lessons • Selecting and adapting activities to develop the 4Cs in CLIL lessons
2	<ul style="list-style-type: none"> • Guiding students in completing tasks in CLIL lessons • Designing lesson plans for CLIL lessons • Evaluating a CLIL lesson
3	<ul style="list-style-type: none"> • Defining professional learning communities (PLCs) and their operation • Introduction to the Lesson Study model
4	<ul style="list-style-type: none"> • Group mentoring: <ul style="list-style-type: none"> - Designing CLIL lesson plans - Planning the implementation of the Lesson Study model
5	<ul style="list-style-type: none"> • Micro-teaching practice
6	<ul style="list-style-type: none"> • Micro-teaching practice
7	<ul style="list-style-type: none"> • Finalizing the implementation plan for the Lesson Study model at the local level • Developing an activity plan for the professional learning community
8	<ul style="list-style-type: none"> • Sharing practices on creating an English-rich learning environment for students • Applying ICT and Artificial Intelligence in activities to create and enhance students' English exposure