

2030 雙語政策下中外師協同教學現況、困境及因應策略之研究：以臺南市國小為例

陳姝澄

臺南市安平區新南國小教師

臺南市教育局新課綱創思教學研發中心教學研究教師

國立臺南大學教育學系課程與教學研究所博士班研究生

中文摘要

在全球化與知識經濟的趨勢下，英語已成為學習與競爭力的關鍵；臺灣自 2018 年推動「2030 雙語國家政策」後，中外師協同教學（FET-LET 協同教學）成為國小英語課室的重要實踐，其品質攸關政策成效。本研究以系統性文獻回顧與整合式綜述為主，並以臺南市 30 組 FET-LET 的不介入式課室觀察為輔，建構「『條件—歷程—結果』（CPO）機制模型」，檢視協同教學在政策脈絡下的運作樣態與致效條件。結果顯示：有效協作仰賴例行性共同備課、明確角色分工，以及「計畫—教學—觀察—回饋」循環；然而多數課堂仍停留於表層協作／指令對齊層級，常見指令可理解度不足與任務—評量不對齊等瓶頸。據此，本研究提出六項短期專業學習模組與「最小可行方案（MVP）：每週至少 20 分鐘受保護之共同備課、角色矩陣（R/A/C/I），以及 2-3 條可見的成功指標」，作為學校與系統層面推動與擴散之可操作槓桿。

Exploring the Effectiveness and Challenges of Co-Teaching under Taiwan's 2030 Bilingual Policy: A Literature Review and Field Observation in Tainan

Shu-Cheng, Chen

Teacher, Sinnan Elementary School, Anping District, Tainan

English Education Researcher and Teacher Trainer, Office of Educational Research and Innovation on the 12-Year Curriculum Guidelines

Doctoral Student, Graduate Institute of Curriculum and Instruction, Department of Education, National University of Tainan, Taiwan

Abstract

Following Taiwan's Bilingual 2030 initiative, FET-LET co-teaching became a core practice in elementary English classrooms, and its quality was pivotal to policy outcomes. This study combined a systematic review and integrative synthesis with nonintervention classroom observations of 30 FET-LET pairs in Tainan and developed a Conditions-Processes-Outcomes (CPO) mechanism model to explain how co-teaching operated under the policy. Findings indicated that effective collaboration hinged on routinized joint planning, explicit role allocation, and plan-teach-observe-feedback cycles. Yet most observed classes remained at surface-level collaboration/directive alignment and frequently exhibited incomprehensible instructional cues and task-assessment misalignment. To address these gaps, we proposed six short professional-learning modules and a minimum viable package (MVP)—at least 20 minutes of protected weekly co-planning, a role matrix (R/A/C/I), and two to three visible success criteria—as practical levers for school- and system-level implementation and scaling.

Keywords: Bilingual 2030; co-teaching; FET-LET; role matrix; task-assessment alignment

1. Introduction

1.1 Background and Rationale

Amid globalization and the knowledge economy, English has become a core vehicle for cross-border communication, academic inquiry, and professional competitiveness; accordingly, many governments treat language education as a strategic lever for national development. Since Taiwan launched the Bilingual 2030 policy in 2018, bilingual education has progressively extended into compulsory schooling with the dual aims of strengthening population-level English proficiency and cultivating intercultural competence. The elementary stage—developmentally sensitive and highly malleable—has thus become a critical site for policy implementation and classroom practice.

To advance the policy, the Ministry of Education has continued to recruit Foreign English Teachers (FETs) and encourage Local English Teachers (LETs) to adopt co-teaching. Theoretically, FETs provide authentic contexts and target-like pronunciation that increase the naturalness of language input, whereas LETs—grounded in curricular familiarity and knowledge of learner diversity—offer scaffolding and cultural bridging. When collaboration is effective, the roles are complementary and jointly promote learners' motivation and achievement. However, empirical work also shows that without protected joint planning time, explicit role allocation, and intercultural communication supports, collaboration tends to become superficial and may trigger disputes over roles and responsibilities, disrupting classroom processes and learning outcomes.

Building on international and domestic scholarship, this study traces a shift from asking “Does it work?” to “How do we govern quality?” Key emphases include routinized joint planning, clear role allocation, directive alignment (prior agreement on what and how task/management directions are delivered), code coordination (planned use of L2 and L1 by function, timing, and proportion), and the use of process indicators—such as the student-student interaction ratio, seconds per speaking turn, and visible success criteria—to monitor classroom quality. Domestically, research under Bilingual 2030 has moved from descriptive accounts to PLC-based improvement cycles that stress task-assessment alignment, speaking-oriented task design, and low-burden data feedback, with growing attention to equity and supportive multimodal input. In response, we complement a systematic review and integrative synthesis with nonintervention observations of 30 FET-LET pairs across 30 elementary schools in Tainan to examine operational status, challenges, and responses

under the policy context. Through dialogue between theory and field observation, we propose an actionable package—at least 20 minutes of protected weekly co-planning, a role matrix (R/A/C/I), and two to three visible success criteria with feedback cycles—to help teams cultivate complementarity rather than counterbalancing and thereby enhance students' English learning.

1.2 Significance of the Study

Theoretical. Prior Taiwan-focused work often centers on policy rollout and teacher narratives, with fewer studies offering cross-study synthesis or reading field evidence against theory. By combining a systematic review with an integrative synthesis, we construct a Conditions-Processes-Outcomes (CPO) model that strengthens theorization and generalizability.

Practical. Although FET-LET co-teaching is now common, insufficient joint planning, blurred roles, and loose procedures remain widespread. We offer directly usable tools—a plan-teach-observe-feedback cycle, a role matrix (R/A/C/I), and a language-scaffolding checklist (sentence frames/visual supports)—to support classroom implementation.

Policy. Beyond expanding “how much” (coverage), the policy must ensure “how well” (quality). Our recommendations focus on time governance, process standards, and capacity building as levers for central and local authorities.

1.3 Objectives and Research Questions

Objectives.

- (1) Clarify the theoretical and policy background of FET-LET co-teaching, with reference to transferable elements from JET/EPIK;
- (2) Analyze, in light of Tainan field observations, how joint planning, role allocation, and code coordination shape classroom processes and learning outputs;
- (3) Identify key antecedents of effectiveness and propose actionable, school-based strategies and procedural guidance.

Research Questions.

RQ1 (literature-led): In elementary EFL/CLIL settings, what are the main models, recurrent benefits, and typical challenges of FET-LET co-teaching (including

transferable elements from JET/EPIK)?

RQ2 (integrative): How do literature-identified antecedents—joint planning, role allocation, intercultural communication, administrative support, and task-assessment alignment—manifest across the 30 FET-LET pairs in Tainan, in terms of frequency and form?

RQ3 (model-building): Combining evidence from the review and the observations, what Conditions-Processes-Outcomes (CPO) mechanism accounts for co-teaching effectiveness, and what school-level strategies are actionable as a result?

1.4 Definitions and Scope

Key terms.

- FET-LET. Foreign English Teacher and Local English Teacher, respectively.
- Co-teaching. FET and LET teach the same class period collaboratively rather than splitting time or subjects. Common models: team teaching (both address the whole class), parallel teaching (split groups, same objective), station teaching (students rotate through tasks), and support/alternative teaching (one leads, one provides targeted support).
- EFL / CLIL. English as a Foreign Language / Content and Language Integrated Learning. The focal setting of this study is elementary EFL; CLIL cases are referenced where relevant.
- Directive alignment. Prior alignment of what directions are given and how they are delivered (content, sequence, language form, modality), so students can understand and begin tasks without line-by-line translation.
- Code coordination. Planned use of L2 and L1 by function, timing, and proportion (e.g., brief L1 bridge → L2 restatement), with explicit signals for switching to preserve comprehensibility and task focus.
- Role matrix (R/A/C/I). A lightweight matrix assigning who is Responsible, Accountable, Consulted, and Informed for key lesson stages and routines.
- Success criteria. Two to three visible, student-friendly indicators of quality for a task or product, used for self/peer/teacher feedback.

- School-based micro-credential. A competency-based, stackable badge awarded on evidence of classroom practice (e.g., completed three-column plan, role matrix in use, observation notes, indicator data), verified by school/district procedures.

Scope and positioning.

- Approach. Literature-led, field-corroborated design: a systematic review and integrative synthesis provide the general framework ($\approx 80\%$ weighting), complemented by nonintervention observations ($\approx 20\%$) of 30 FET-LET pairs in Tainan to map local patterns against the framework.
- Boundaries. No experimental intervention or standardized achievement tests were conducted; observations and artifacts are used descriptively to illuminate conditions, processes, and outcomes (CPO) and to inform actionable school-level strategies.
- Setting. Elementary EFL classrooms in Tainan. Where CLIL lessons are encountered, analysis foregrounds language-learning processes and task-assessment alignment.
- Academic-year convention. “Academic year” refers to August 1-July 31, which also anchors timelines and scheduling references in the study.

2. Literature Review

2.1 Policy Architecture and International Governance Lessons

Since 2018, Taiwan’s Bilingual 2030 policy has framed co-teaching as part of a tripartite governance agenda—people (staffing and capacity), curriculum (standards and materials), and settings (contexts and resources)—with FET-LET collaboration increasingly institutionalized in elementary EFL classrooms (Ministry of Education, 2018; National Development Council, 2022; Lin & Wu, 2025). International large-scale programs provide transferable design elements. Japan’s JET and Korea’s EPIK emphasize clear role division, intercultural training, routinized joint planning, and monitorable process/outcome indicators (e.g., directive alignment, code coordination, student-student interaction ratio, seconds per speaking turn, visible success criteria), forming a multilevel governance chain from central authorities to schools and classrooms (Shin et al., 2021; Kim & Moodie, 2023). Taiwan’s direction is broadly consonant with these trends, but implementation still requires time governance, process standards, and capability building to avoid superficial collaboration.

2.2 Co-Teaching in Chinese EFL/CLIL Contexts (2007-2025): Emergence → Quality Governance → Policy Coupling

To trace development in Chinese EFL/CLIL settings, the literature of roughly two decades can be grouped into three phases.

Phase I: Emergence and Problematization (2007-2014).

Predominantly qualitative case and field studies documented gaps between policy and classroom realities: blurred roles, absent co-planning, and governance frictions. Findings cautioned that, without structural supports, co-teaching often regresses to “two teachers present, one teacher instructing,” yielding limited learning gains (Luo, 2007; Chen, 2010; Park, 2014).

Phase II: Quality Governance (2015-2022).

The focus shifted from whether co-teaching works to how to make it work. Core levers included protected joint planning, explicit role allocation, and task-assessment alignment, typically embedded in professional learning communities (PLCs). Evidence consistently showed gains in listening/speaking and motivation, while transfer to reading/writing depended on task design, scaffolding, and time allocation. School-level governance (e.g., protected co-planning windows, resource allocation) and classroom routines (directive alignment, code coordination) emerged as mediating conditions (Beatty, 2019; Lee, 2021; Shin et al., 2021; Graham & Yeh, 2022; Chien, 2022).

Phase III: Model Optimization and Policy Coupling (2023-2025).

Recent work proposes a collaboration spectrum—from L1 surface-level collaboration to L4 co-construction—and routinized plan-teach-observe-feedback cycles, with context-sensitive selection among team/parallel/station/support models (Kim & Moodie, 2023). Concurrently, studies advocate low-burden, data-informed indicators (e.g., student-student interaction ratio, seconds per speaking turn, visible success criteria) and attention to equity (benefits for disadvantaged and struggling learners) to monitor quality and outcomes (Anggrainy et al., 2024; Cvitkovic, 2025). These trajectories align with Taiwan’s current implementation needs, shifting the emphasis from staffing to governing quality.

2.3 Mapping Conditions-Processes-Outcomes (CPO)

Conditions. Quality of joint planning; a role matrix (R/A/C/I); intercultural

communication supports; and administrative time/resources.

Processes. Directive alignment of task/management directions; code coordination (planned L2/L1 use by function, timing, and proportion); question levels - lower-order thinking (LOT; recall/identify) through higher-order thinking (HOT; explain/compare/justify); student-student interaction; and feedback cycles.

Outcomes. Comprehensibility and participation leading to task performance under task-assessment alignment (e.g., product quality, seconds per speaking turn, completion rates). Collectively, the literature points to a mechanistic linkage from conditions through processes to outcomes rather than effects driven by any single activity (Park, 2014; Lee, 2021; Graham & Yeh, 2022). See Appendix A (glossary) and Appendix B (task structures/tools).

2.4 Learning Effects and Boundaries

Most studies report positive effects on listening/speaking, pragmatics, and motivation under co-teaching, especially for output-oriented, task-based lessons with visible success criteria. Extension to reading/writing requires staircase scaffolding (lexis → sentence → paragraph) and deliberate time allocation; absent standard routines and role clarity, classrooms tend to revert to co-presence without collaboration, lowering interaction quality (Beatty, 2019; Graham & Yeh, 2022; Lee, 2021). Two actionable anchors repeatedly recommended are success-criteria-task-assessment alignment and an oral-task repertoire that normalizes student output.

2.5 Gaps and Present Study Positioning

Integration gap. A unifying framework that connects collaboration spectrum, role matrix, and plan-teach-observe-feedback with CPO is still uncommon, limiting transfer across contexts.

Methods gap. Limited in-situ observation and longitudinal tracking constrain evidence on micro-interaction, role negotiation, and real-time decision-making (heavy reliance on self-report).

Outcomes gap. Shortage of quantitative, learning-evidence indicators (e.g., oral fluency, task completion, participation traces) and equity analyses.

Governance gap. Inputs (time, training) are not consistently linked to process and outcome indicators for comparability.

Positioning of this study. We address these gaps via a systematic review and integrative synthesis anchored in CPO, complemented by nonintervention observations of 30 FET-LET pairs in Tainan to map local patterns against the framework, and we propose minimum-viable implementation levers (protected weekly co-planning, role matrix R/A/C/I, two to three success criteria) for school-level uptake and scaling.

3. Methods

3.1 Design and Approach

This study examines the enabling conditions and breakdown mechanisms of FET-LET co-teaching under Taiwan's Bilingual 2030 policy. Given substantial school-level heterogeneity, the absence of common indicators, and our nonintervention stance (no instructional treatments or standardized outcome tests), we adopted an integrated approach: evidence synthesis (literature-led) plus field corroboration (nonintervention observation).

- Primary method: systematic review and integrative synthesis to build a Conditions-Processes-Outcomes (CPO) conceptual model, with quality-weighted evidence.
- Supplementary method: multi-site, nonintervention classroom observations documenting interaction and collaboration patterns for 30 FET-LET pairs across 30 public elementary schools in Tainan, used as contextual exemplars and checks.
- Integration logic: literature provides generalizable trends and a theoretical scaffold; observations supply visible behaviors and situated context. We triangulate via model benchmarking and case mapping. Overall reporting weight \approx literature 80% / observation 20%.

3.2 Operational Definitions and Conceptual Framework

Co-teaching. FET and LET collaborate within the same instructional period in one of four models: team, parallel, station, or support/one teach-one assist; this excludes mere time-sharing or subject splitting.

Collaboration spectrum (L1-L4).

- L1 Surface-level collaboration: ad-hoc division of labor, no co-planning, divergent goals.
- L2 Directive alignment: synchronized procedures and signals, but limited

integration of scaffolds/assessment.

- L3 Process collaboration: routinized joint planning, explicit role allocation, operational feedback cycles.
- L4 Co-construction: jointly designed goals-activities-assessment, PLC-like shared responsibility.

Effectiveness indicators.

- Process: student participation; T-Ss vs. Ss-Ss interaction ratio; task completion; question levels (LOT → HOT); code coordination (planned L2/L1 functions, timing, and proportion; visible switch signals).
- Output: seconds per speaking turn, use of language chunks/frames, and task product quality. (No standardized tests were administered.)

CPO linkage.

- Conditions: time governance (protected joint planning), tool governance (role matrix [R/A/C/I], three-column co-planning sheet [Goals-Activities-Assessment], two to three visible success criteria), professional capability (comprehensible input, multimodal directions, code coordination), and administrative resources.
- Processes: plan-teach-observe-feedback cycles that enact task-assessment alignment and segmental co-leadership.
- Outcomes: increased participation and oral output, stable task completion, and downstream reading/writing carryover.

3.3 Systematic and Integrative Review Methods

Scope and sources. 2005-2025; English and Chinese; databases: Web of Science, Scopus, ERIC, ProQuest Dissertations, Google Scholar; gray literature includes MOE/NDC documents. Phasing for synthesis: 2007-2014 (emergence/problematisation), 2015-2022 (quality governance), 2023-2025 (model optimization/policy coupling).

Search strategy. Co-teaching/team-teaching, FET/LET, bilingual policy, EFL/CLIL, elementary/primary, joint planning, role allocation/division, intercultural communication, JET, EPIK (Boolean AND/OR; synonym expansion). Chinese keywords paralleled these terms.

Inclusion/exclusion. Include EFL/CLIL studies (elementary/junior-high) on FET-LET collaboration; qualitative/quantitative/mixed methods and program evaluations focusing on co-planning, role division, interaction, code coordination, administrative support, and task-assessment alignment. Exclude theory-only pieces lacking context, single-teacher accounts without collaboration, non-retrievable full texts, or low-credibility reports.

Quality appraisal and extraction. We had applied Critical Appraisal Skills Programme (CASP)- and Mixed Methods Appraisal Tool (MMAT)-informed criteria (design clarity, sampling transparency, tool validity/reliability, analytic adequacy) to tag studies as high, medium, or low quality; we had logged a PRISMA-style selection trail (Page et al., 2021); and we had extracted site, sample, method, findings, and quality tags into a master matrix.

Synthesis. Thematic aggregation and narrative synthesis with evidence weighting (prioritizing medium/high-quality studies), following PRISMA for reporting and CASP/MMAT for appraisal (see Methods, Section 3.3); item-level checklists are not reproduced.

3.4 Multi-Site Nonintervention Observations

Purpose. This component provides corroborative, context-grounded evidence to benchmark and interpret the systematic review; analyses are descriptive and non-causal.

Sites and participants. 30 public elementary schools in Tainan; one FET-LET pair per school (N=30). Grades: primarily middle to upper elementary (per scheduling).

Rationale for site selection. Tainan was chosen because the author, as an assigned non-intervention observer, could lawfully and ethically conduct low-intrusion classroom observations; the city has maintained diverse FET-LET implementations since 2018, enabling close alignment with the study's CPO mechanism; and deriving actionable levers (MVP: institutionalized weekly co-planning, role matrix, visible success criteria) benefits from fine-grained description within a single, coherent administrative context. This design supports analytic rather than statistical generalization.

Dosage and waves. Each pair observed 1-2 lessons (total ≥ 30 -60 lessons), in two waves:

Wave 1: 15 pairs, Sep 2023-June 2024.

Wave 2: 15 pairs, Sep 2024-June 2025.

Researcher role. Nonparticipatory observer; no scoring; no disruption to instruction.

Instruments and records.

- Structured observation sheet (short form): grade, period, topic; co-teaching model; interaction distribution (T-Ss, Ss-Ss; 30-60-s sampling); code coordination (L2/L1 switch points and rules); time structure (engage-input-practice-output shares); co-planning traces (three-column plan, role matrix, co-produced materials); division of labor and real-time coordination (segmental leadership, hand/visual signals).
- Post-lesson semi-structured memo (criterion-based sampling): To minimize intrusion, only 10 pairs—those showing L1-L2 levels and multiple risk indicators (sparse co-planning trace, task-assessment misalignment, low comprehensibility of directions)—received in-situ semi-structured conversations (\approx 60-120 minutes) to clarify constraints, communication context, and real-time decisions. Other pairs: observation notes only.
- We did not conduct student interviews or record student-identifiable information. Incidental board/projected visuals (no faces) were photographed when permitted to document flow.

Procedures. Consent from schools and teachers \rightarrow scheduled in-class observation (real-time notes; non-person images when allowed) \rightarrow after Wave-1 initial passes, select 10 pairs for on-site conversations using pre-set criteria \rightarrow same-day de-identification of summaries and memos \rightarrow thematic coding.

Data handling. Descriptive statistics (means/ranges) for quantifiable items (e.g., Ss-Ss ratio, segment timing, L2/L1 switches, seconds per speaking turn as a class-level estimate). Selected de-identified vignettes illustrate typical high/low-collaboration segments. These are cross-sectional descriptive data; no causal inference or city-level generalization is claimed.

3.5 Literature-Observation Integration Strategy

1. Benchmarking: use the synthesized collaboration spectrum \times conditions \times effectiveness as the yardstick.

2. Case mapping: position each class (e.g., “L2 directive alignment and limited co-

planning and low Ss-Ss”) on that yardstick.

3.Ga diagnosis: compare “literature-recommended” vs. “field-observed” to identify systemic (time/resources) and operational (scaffolds/process) contributors.

4.Strategy extraction: derive school-ready levers (≥ 20 -minute protected weekly co-planning; role matrix [R/A/C/I] card; two to three success criteria mini-rubrics; monthly plan-teach-observe-feedback cycle).

3.6 Trustworthiness and Bias Mitigation

Multiple data sources; quality-weighted synthesis; standardized observation tools; peer calibration of coding; thick description via vignettes; researcher reflexivity to separate literature claims from field impressions and to avoid over-generalizing from cases.

3.7 Ethics

We did not interview students, nor did we collect identifiable student data; observations targeted whole-class dynamics and process quality.

Because the project also supports policy/program evaluation in partnership with authorities, we implemented two-track data governance:

- Policy/program evaluation track (MOE/LEA): teacher-identifiable observation records may be retained within controlled administrative systems only for aggregated policy/program evaluation and management; no transfer of identifiable records to the research track.
- Academic research track (researchers): analyses and publications use de-identified/anonymous data only; no personnel evaluation or HR use. Research data are encrypted and retained up to five years (or deleted earlier when the purpose is fulfilled); any new purpose requires additional notice and consent. If incidental student images/audio occur, they are promptly masked or deleted.

3.8 Limitations

Design: no random assignment, no intervention, and no standardized achievement tests → no causal claims; the design is evidence synthesis and contextual corroboration.

Observation: 1-2 cross-sectional visits per class may under-estimate maturation/trajectory effects.

Transferability: cross-country system differences (e.g., JET/EPIK vs. Taiwan) limit portability; the proposed minimum viable package (MVP) aims to reduce contextual dependency.

Quantification: emphasis on process and low-inference output indicators; future work should add validated oral-fluency measures.

Interview bias: criterion sampling focused on weaker patterns and relatively long conversations (60-120 minutes) may inflate negative proportions and recall bias; we flag this descriptively in the Results/Discussion.

3.9 Timeline and Outputs

School year definition: Aug 1-Jul 31.

Observation waves: Wave 1 = 15 pairs (Sep 2023-June 2024); Wave 2 = 15 pairs (Sep 2024-June 2025).

Overall schedule.

- Aug-Nov 2023: searching, quality appraisal, thematic synthesis; draft CPO model; Wave-1 observations in parallel.
- Sep 2023-June 2024: complete Wave 1; rolling descriptive stats and vignettes.
- Feb-Jul 2024: integrated Wave 1 interim findings; calibrate tools/indicators.
- Sep 2024-June 2025: complete Wave 2; compile descriptive stats and vignettes.
- Feb-Jul 2025: finalize literature-observation crosswalk; extract strategies; prepare appendices (co-planning checklist, role matrix [R/A/C/I], language scaffolds, lesson-flow cards).

3.10 Summary

Chapter 3 defines the methodological stance and integration pathway: a systematic review and integrative synthesis to establish the co-teaching effectiveness mechanism (CPO), corroborated by two-wave nonintervention observations of 30 FET-LET pairs in Tainan. This enables Chapter 4 to report findings through a

consistent “literature trends → local patterns → aligned strategies” arc, and positions Chapter 5 to translate results into actionable, monitorable, and scalable school and policy recommendations.

4. Results and Discussion

This chapter reports findings primarily from the systematic review and integrative synthesis, supplemented by nonintervention observations of 30 FET-LET pairs in Tainan (approximate evidence weighting: 80% literature / 20% observations).

4.1 Thematic Findings from the Literature

4.1.1 Collaboration Models and Structural Requirements

Across international and Taiwan-focused studies, four co-teaching modalities recur: team teaching (both teachers address the whole class), parallel teaching (split groups taught simultaneously), station teaching (students rotate through task stations), and support/one-teach-one-assist (one leads, one provides targeted support). High-quality collaboration consistently rests on three structural requirements:

1. Routinized joint planning that aligns goals-activities-assessment (not just procedural coordination);
2. Explicit role allocation captured in a role matrix (R/A/C/I) that assigns who is Responsible, Accountable, Consulted, and Informed; and
3. A continuous plan-teach-observe-feedback cycle.

To support terminological precision and classroom uptake, key terms are consolidated in Appendix A, and frequently used communicative structures (e.g., Timed Pair Share, Stand-N-Share) appear in Appendix B.

4.1.2 Learning Benefits and Boundaries

Evidence is strongest for gains in listening-speaking performance, pragmatic use, and motivation, especially when tasks are product-oriented and success criteria are made visible. Extending impact to reading/writing depends on time allocation and cross-skill scaffolding (e.g., lexical → sentence → paragraph progression) that is co-planned and assessment-aligned. Where planning is thin or roles are ambiguous, benefits are diluted by pacing pressure and classroom-management overhead.

4.1.3 Recurrent Challenges (Methods → Culture → Governance)

Bottlenecks cluster at three layers:

- Method level: insufficient protected co-planning, task-assessment misalignment, LOT-heavy questioning, and low S-S (student-student) interaction;
- Culture level: divergent teaching beliefs, inconsistent code coordination principles, and asymmetric communication;
- Governance level: weak administrative supports and lack of SOPs or targeted training.

These yield surface-level collaboration/directive alignment (L1-L2) rather than sustained co-construction.

4.1.4 Enablers and Transferable Principles

Improvement hinges on time governance (e.g., ≥ 20 minutes of protected weekly co-planning), tool governance (role matrix; three-column co-planning sheet for goals-activities-assessment; visible success criteria), PLC routines, and data-informed monitoring using process (e.g., S-S ratio, seconds per speaking turn) and output indicators (e.g., task completion). Operational definitions and quick-measure guidance are provided in Appendix B.

4.2 Descriptive Findings from Tainan Observations

Note. This section presents nonintervention, cross-sectional descriptive patterns from 30 FET-LET pairs (30-60 lessons observed across two waves). It provides local context and typical configurations; no causal claims are made. Operational definitions for quick indicators (e.g., S-S ratio, seconds per speaking turn) appear in Appendix B.

4.2.1 Overall Collaboration Levels and Modalities

We begin with an overall impression, then add indicator-based results and short vignettes.

- Most classes clustered at L1-L2 (surface collaboration/directive alignment): procedures were generally synchronized and classroom signals consistent, but co-planning traces were thin, and assessment/scaffolding were weakly linked to tasks.

- A minority reached L3 (process collaboration): regular co-planning, co-produced materials, and visible plan-teach-observe-feedback cycles.
- A very small number approached L4 (co-construction): jointly designed goals-activities-assessment with in-class alternation of leadership and mature in-the-moment coordination.

4.2.2 Classroom Interaction and Code Coordination

Indicator-focused observations emphasized T-S vs. S-S time, seconds per speaking turn, and code-switching principles (see Appendix B).

- Interaction was T-S dominant; S-S time was relatively limited. When structured cooperative routines (e.g., Timed Pair Share, RallyCoach) were used, S-S quality and quantity improved.
- A common division was FET in English with LET providing L1 scaffolds. In a subset of lessons, frequent, unsignaled switches reduced coherence.
- Questioning skewed LOT, with few HOT prompts. Adoption of the speaking-led task set in Appendix B is recommended.

4.2.3 Co-Planning and Role Allocation

Pre-lesson coordination was largely ad hoc, with few cases of a protected co-planning slot. Co-planning artifacts—three-column (goals-activities-assessment) plans, co-authored materials, and a role matrix (R/A/C/I)—were uncommon. The FET typically led interaction, while the LET handled classroom management and L1 scaffolding; translation was mostly reactive. At L3, pairs pre-assigned segmental leadership by phase and used unobtrusive handoff cues; materials were co-produced with visible success criteria and language frames. A small minority neared L4, co-designing objectives-tasks-assessment, mapping responsibilities via R/A/C/I, and keeping brief plan-teach-observe-feedback notes.

4.2.4 Task Design and Learning Outputs

“Surface tasks” (choral reading, short IRE sequences) predominated. When tasks combined authentic context and product (e.g., posters, short oral briefs, interview pairings), speaking turns lengthened and participation broadened. Simple, task-tied success criteria (2-3 items) were not widely used; higher-collaboration classes made

criteria visible and reviewed them at closure.

4.2.5 Vignettes (De-identified)

- Vignette A — Directive alignment (L2): FET leads Q&A → LET translates and boards key items → brief practice; S-S limited. Strengths: clear flow, rapid feedback. Limitations: shallow output.
- Vignette B — Process collaboration (L3): Segmental leadership anchored by a role matrix (R/A/C/I); framing and Timed Pair Share for practice and Stand-N-Share for product. Strengths: broad participation, natural chunk use. Need: sustained co-planning time.
- Vignette C — Near co-construction (L3 → L4): Community-theme task; co-planning yielded chunks-frames-success criteria tables; post-lesson feedback note captured adjustments. Students produced 20-30-second speaking turns. Bottleneck: weaker downstream writing carryover.

4.2.6 Key Deficits and Risk Signals (Integrative Summary)

Recurring deficits

- 1.Capability/PLC gaps: limited pre-service/in-service training on co-planning, role allocation, code coordination, and task-assessment alignment; PLC routines not institutionalized.
- 2.Time/tools not secured: co-planning not protected; underuse of three-column sheets and role matrices; feedback cycles sporadic.
- 3.Comprehensibility/code issues: instructions and input not consistently CI; code-switching lacks shared rules.
- 4.Alignment/assessment weak: success criteria not made visible; task-assessment misaligned, constraining self/peer feedback.
- 5.Curricular/interaction constraints: limited longitudinal coherence; low S-S and short speaking turns.

Plausible mechanism chain (descriptive): training gap → weak co-planning/tools → low instructional comprehensibility → task-assessment misalignment → low S-S/short turns → constrained benefits.

Risk signals (see Appendix B for quick-measure guidance): persistently low S-S; no growth in seconds per speaking turn; no co-planning artifacts in evidence; success criteria not visible; inconsistent code-switching.

Implications for action: In brief, we recommend weekly ≥ 20 -minute protected co-planning, a role matrix (R/A/C/I), and 2-3 visible success criteria, supported by a monthly plan-teach-observe-feedback cycle.

4.3 Literature-Observation Crosswalk

Purpose. This crosswalk aligns literature-derived expectations (CPO model) with patterns observed in 30 FET-LET classrooms in Tainan, highlighting convergences, gaps, and actionable levers.

4.3.1 Convergences

Authentic contexts, visible scaffolding, and an explicit product were consistently associated with stronger oral output. Tasks that coupled situational framing with a tangible product (e.g., poster, brief presentation, peer interview) yielded longer speaking turns and broader student participation, consistent with prior research.

Protected co-planning predicts higher collaboration levels. Pairs that scheduled and documented co-planning (three-column sheets; shared materials) more often operated at L3 (process collaboration) with visible plan-teach-observe-feedback cycles.

Role clarity supports in-class alternation of leadership. Use of a role matrix (R/A/C/I) was associated with smoother segmental leadership (lead/support by phase), echoing findings that clarity of responsibility increases interactional quality.

Structured routines raise S-S. Adoption of cooperative structures (e.g., Timed Pair Share, Stand-N-Share) elevated student-student interaction (S-S) ratio and reduced teacher-dominant talk.

Visible success criteria facilitate feedback. Classes that displayed 2-3 task-tied success criteria used them at closure for feedback/self-monitoring, reflecting literature on assessment for learning.

4.3.2 Divergences (Where the field lags)

Training/PLC gap. Few pairs reported targeted training in co-planning, role allocation, code coordination, or task-assessment alignment; PLC routines were not

institutionalized, limiting movement beyond L1-L2.

Comprehensibility and code coordination inconsistent. Core instructions did not consistently meet comprehensible input (CI) principles; L2/L1 switches were sometimes frequent and unsignaled, reducing coherence.

Task-assessment misalignment. Success criteria were often absent or generic; assessment artifacts did not map clearly to task demands, weakening feedback and learner self-regulation.

Thin co-planning artifacts. Many lessons showed only ad hoc pre-coordination (verbal/chat), with limited evidence of three-column planning or co-produced materials.

Interaction profile skewed T-S; HOT prompts rare. S-S time and seconds per speaking turn remained modest in many classes; questioning stayed at LOT, with few HOT moves.

4.3.3 Process Breakpoints (CPO lens)

Cross-reading suggests recurring breakpoints in the Conditions → Processes → Outcomes chain:

1. Conditions not secured: limited protected time; missing tools (role matrix, three-column sheet); scant capability building.

2. Process slippage: directive alignment not rehearsed; code switching not pre-agreed; success criteria not made visible; feedback cycle irregular.

3. Outcome suppression: low S-S, short speaking turns, and uneven task completion—especially when product expectations were implicit or criteria were absent.

A descriptive mechanism chain is as follows: training gaps weaken co-planning and the use of shared tools, which lowers instructional comprehensibility; this, in turn, produces task-assessment misalignment and depresses student-student (S-S) interaction and speaking-turn length, thereby constraining learning gains.

4.3.4 Priority Gaps and Implementation Levers

Priority gaps: (a) co-planning protection and artifacts; (b) explicit role allocation; (c) CI-grounded directive alignment and code coordination; (d) task-assessment

alignment with 2-3 visible success criteria; (e) routine plan-teach-observe-feedback cycles.

- Levers for uptake:
 - We propose a minimum viable package (MVP) comprising at least 20 minutes of protected weekly co-planning, a role matrix (R/A/C/I), and two to three visible success criteria.
 - Six short capability modules (90-120 min each): co-planning scripts; CI and multimodality; collaborative teaching scripts; task-assessment alignment; code coordination & intercultural communication; PLC routines.

Quick indicators (see Appendix B): S-S ratio, seconds per speaking turn, task completion rate—tracked monthly to guide PLC reflection.

Summary. The crosswalk shows strong alignment on what works (authentic/productive tasks, protected co-planning, role clarity, visible criteria) and pinpoints why many classrooms remain at L1-L2 (training/time/tools deficits leading to process breakpoints). Addressing the priority gaps via the MVP and short modules is a feasible pathway to scale L3 practices and seed L4 co-construction.

5. Conclusion and Recommendations

5.1 Conclusions

This study adopted a systematic review and an integrative synthesis, complemented by nonintervention observations of 30 FET-LET pairs in Tainan, to address RQ1-RQ3. Three conclusions follow:

1. A three-pillar mechanism governs effective co-teaching. High-quality FET-LET collaboration consistently rests on (a) routinized joint planning with task-assessment alignment, (b) explicit role allocation (role matrix, R/A/C/I), and (c) plan-teach-observe-feedback cycles. When any pillar is weak, classes typically remain at collaboration levels L1-L2 (surface collaboration/directive alignment), yielding limited, hard-to-observe learning outputs.

2. Observable gains concentrate in speaking and pragmatic use. When tasks are output-oriented and supported by visible scaffolds (language chunks/frames) and two to three success criteria, student participation, number of dialogic turns, and seconds per speaking turn increase.

3. Governance is the proximal enabler. Schools that protect weekly co-planning time, operate PLC routines, and use standard tools (three-column co-planning sheet, role matrix, success-criteria mini-rubrics) are more likely to reach L3 process collaboration. Where co-teaching capacities (directive alignment, code coordination, Comprehensible Input and multimodality) are underdeveloped, we observed low instructional comprehensibility, task-assessment misalignment, and fragmented lesson flow.

5.2 Answers by Research Question

RQ1 (Models, benefits, challenges). Four workable models (team, parallel, station, support) each have context-fit conditions. Benefits concentrate in speaking/motivation; recurrent challenges span methods (insufficient co-planning; misaligned tasks and assessments), culture (divergent beliefs; inconsistent code coordination), and governance (time and SOP deficits).

RQ2 (Local presentation of antecedents). In Tainan, most pairs fell at L1-L2, with ad hoc division of labor, unclear/over-translated directions, and missing or invisible success criteria. A minority at L3 showed co-developed materials, segmental leadership, and documented feedback, with more stable outputs.

RQ3 (CPO model and strategies). We propose a Conditions-Processes-Outcomes (CPO) mechanism: governance and capacity inputs (time, SOP, capability modules) enable plan-teach-observe-feedback processes, which in turn increase task-assessment alignment, code coordination, and S-S interaction, leading to improved participation and speaking outputs. A minimum viable package (MVP)—at least 20 minutes of protected weekly co-planning, a role matrix (R/A/C/I), and two to three visible success criteria—offers a practical lever for rapid improvement.

5.3 Recommendations

National level (MoE/K-12 Administration)

1. Define FET-LET collaboration competencies and micro-credentials. Core modules: joint-planning scripts, Comprehensible Input (CI) and multimodality, collaborative teaching scripts, task-assessment alignment, code coordination/intercultural communication, and PLC routines. Issue online modules and school-based micro-credentials with annual review.

2. Formalize time and observation within accountability systems by mandating at

least 20 minutes of protected weekly co-planning and a minimum of one bidirectional observation-feedback cycle per month, and embedding both requirements in teacher professional development and school evaluation metrics.

3. Build a data-informed dashboard that tracks S-S interaction share, seconds per speaking turn, task-completion rate, success-criteria visibility, co-planning time, and cycle-completion rate, and provide school-level feedback without cross-school league-table comparisons.

4. Issue national SOPs for onboarding-co-planning-observation. Align pre-service, onboarding, and in-service supports with clear directive alignment and code-coordination agreements before term start.

Local level (City/County Bureaus)

1. PLC hub schools. Host open plan-observe-feedback days showcasing three-column plans, role matrices, language chunks/frames, and feedback records.

2. Mobile capacity-building teams. Deliver the six modules on site; pair workshops with in-class coaching to ensure enactment.

3. Localized code-coordination guides. Provide bilingual classroom language banks and visual direction cards to raise instructional comprehensibility.

School level (Operational pathway)

1. One-year blueprint (aligned to the Taiwan school year).

Start-up (Aug-Sep): Form the co-teaching team; schedule protected weekly co-planning; adopt the three-page co-planning pack (three-column sheet, role matrix, success-criteria mini-rubric).

Pilot (Oct-Dec): Run a monthly plan-observe-feedback cycle; integrate at least two speaking-led tasks; collect baseline S-S and speaking-turn seconds.

Scale (Feb-Apr): Expand to grade-level teams; start a task-assessment alignment repository; operate the dashboard.

Institutionalize (May-Jul): Finalize school SOPs; hold open lessons; compile data for next-year improvement.

2. SOP essentials. Co-plan 20 minutes (three columns; observable indicators per

column); segmental leadership matrix (who leads/supports by phase and language function); success-criteria display (2-3 per lesson, revisited at closure); quick feedback (oral and mini-rubric; feed the dashboard).

Teacher development (FET/LET)

Six-module monthly cycle paired with plan-observe-feedback practice.

Routine CI and multimodal supports (TPR, visuals, stepwise modeling, sentence frames).

Speaking-led task ladder (self-intro → describe → compare → explain reasons → short opinions) with aligned chunks/frames and success criteria.

Code-coordination pact: English for core communication; Mandarin for scaffolding/safety, with pre-planned switch points.

From LOT to HOT: At least one reason/choice/compare prompt per lesson, with explicit frames (e.g., I think..., because...).

5.4 Monitoring and Evaluation

Core monthly indicators:

S-S interaction share $\geq 35\%$ (target 45% over time)

Seconds per speaking turn ≥ 20 (typical 15-30 range)

Task-completion rate $\geq 85\%$

Success-criteria visibility $\geq 90\%$

Co-planning time achievement $\geq 80\%$

Plan-teach-observe-feedback cycle completion $\geq 80\%$

Supplementary: ≤ 1 unclear-direction incident per lesson; ≥ 1 co-created material per unit; PLC participation $\geq 80\%$.

5.5 Limitations

This study offers evidence synthesis and nonintervention observation rather than causal inference. Each pair was observed 1-2 times, which may understate longitudinal growth; cross-national transferability is bounded by institutional differences; outcomes

rely on process-proximal indicators (speaking turns, participation), not standardized achievement tests.

Setting. The sample is limited to one city-Tainan. The study advances analytic generalization within a bounded governance context and does not claim statistical generalization.

Unmeasured structural covariates. Teacher L2 proficiency, professional-development intensity, and contract stability were not directly measured; these unobserved structural factors may confound process indicators and should be modeled in subsequent work.

5.6 Future Research

Cross-county matched comparison. Conduct matched comparisons across counties (urban/rural, school size, FET tenure/turnover) with common quick indicators—student-student (S-S) interaction share, seconds per speaking turn, task-completion rate, and success-criteria visibility—and shared SOPs; analyze with multilevel models to test the stability of CPO pathways across contexts.

Longitudinal and quasi-experimental extensions of the MVP. Implement at least a one-year follow-up with low-intensity quasi-experiments in volunteer schools. Use fluency (words per minute, pause ratios), comprehensibility, and task-completion as primary outcomes, and include teacher-level covariates (CEFR self-ratings or standardized proficiency bands; PD hours and micro-credential accrual; FET/LET tenure and turnover).

Downstream transfer. Examine reading/writing carryover from speaking-led tasks under aligned scaffolding.

Equity analyses. Investigate heterogeneous effects for disadvantaged or struggling learners.

Micro-interaction studies. Analyze code-switching timing, gesture use, and visual supports as micro-scaffolds linked to task performance.

Cost-effectiveness. Model the relationship between investments (time, equipment) and both process and outcome indicators.

5.7 Closing Statement

Under Taiwan's Bilingual 2030 policy, the pivotal question is not whether

external personnel are present, but whether collaboration is governable. By coupling time governance, tool governance, and capacity modules, and by deploying a pragmatic MVP, schools can move most classes from L1-L2 to L3, and progressively toward L4 co-construction—enabling students to learn in authentic, comprehensible, and assessable classrooms and to accumulate visible gains in English proficiency and confidence.

Appendix A. Glossary and Abbreviations

FET-LET: Foreign English Teacher / Local English Teacher.

FET-LET co-teaching: Two teachers jointly teach one class using team, parallel, station, or support (one teach-one assist) modes.

CPO model: Conditions-Processes-Outcomes mechanism.

Directive alignment: Pre-agreed what/how/when/in which language directions are delivered so students can act without line-by-line translation.

Code coordination: Planned L2/L1 function, timing, and proportion, with clear switch signals.

Role matrix (R/A/C/I): Responsible, Accountable, Consulted, Informed role clarity.

PLC: Professional Learning Community; routine plan-teach-observe-feedback cycle.

Success criteria: 2-3 visible, student-friendly statements defining “what good looks like.”

S-S interaction share: Proportion of time students talk with students.

Seconds per speaking turn: Average duration of a student oral turn during task time.

LOT / HOT: Lower-order vs. higher-order thinking (recall/identify → explain/compare/justify).

CI and multimodality: Comprehensible Input supported by visuals, gestures, modeling, and language chunks.

Task-assessment alignment: Task, success criteria, and quick rubric point in the

same direction.

Segmental leadership: Alternating lead/support by phase (Engage → Practice → Product).

MVP: Minimum viable package—at least 20 minutes of protected weekly co-planning, a role matrix (R/A/C/I), and 2-3 visible success criteria.

Appendix B. Co-Planning Mini-Kit and Monitoring Indicators (one page)

B1. 20-Minute Protected Co-Planning (three-column: Goals-Activities-Assessment)

(1)Goals: target language/function and LOT → HOT prompt.

(2)Activities: structure (e.g., Timed Pair Share, Stand-N-Share), timing, materials.

(3)Assessment: 2-3 success criteria and a quick 0-2 micro-rubric. Also fix: code-coordination plan, roles (R/A/C/I), likely risks (time, clarity, noise).

Micro-rubric (example)

Criterion (postcard talk)	0 = not yet	1 = emerging	2 = met
20-30-sec talk	<15 s or stops	15-19 s	20-30 s
Two reasons with <i>because</i>	none/1	2 with prompts	2 independently
Audibility/eye contact	seldom clear	mostly clear	clear & steady

B2. Role Matrix (R/A/C/I) by Lesson Phase

Phase	L e a d S u p p o r t I n f o r m			Notes
	(R/A)	(C)	(I)	
Engage / model	FET	LET	Admin	L2 model + visuals/TPR
Guided practice	LET	FET	—	Checks for understanding
Product / share	FET	LET	—	Criteria visible on screen
Management / safety	LET	FET	Admin	L1 for safety/clarity

(Swap by context; update per unit.)

B3. “Quick Clarity” Card (directive alignment & CI)

One concise on-screen direction and live demo and gesture; You-do/We-do/Pair-do; optional ≤10-15-sec L1 bridge, then restate in L2.

B4. Monitoring Indicators (definitions & targets)

- S-S interaction share = student-to-student talk ÷ task time. Target: $\geq 35\%$ (toward 45%).
- Seconds per speaking turn = mean turn length from 60-120-sec spot checks. Target: ≥ 20 s (typ. 15-30 s).
- Task-completion rate = % producing intended product on time. Target: $\geq 85\%$.
- Success-criteria visibility rate = lessons with criteria posted/referenced. Target: $\geq 90\%$.
- Co-planning hours met = weeks with ≥ 20 -min protected slot. Target: $\geq 80\%$ of weeks.
- Cycle completion rate = months with ≥ 1 plan-teach-observe-feedback loop. Target: $\geq 80\%$.
- Auxiliary guardrail: incomprehensible FET directive incidents ≤ 1 per lesson (spot-check notes).

References

- Anggrainy, S., Matsumiya, N., & Watanabe, T. (2024). A case study of Japanese teachers' self-reflection on elementary English lessons. *SN Social Sciences*, 4(11), Article 207. Retrieved from <https://doi.org/10.1007/s43545-024-01002-8>
- Beatty, A. (2019). A phenomenological study of co-teaching' s collaborative decision-making process in Taiwanese elementary school [Master' s thesis, St. Cloud State University]. *St. Cloud State University Repository*. Retrieved from https://repository.stcloudstate.edu/tesl_etds/14/
- Chen, C.-W. Y. (2010). A case study of non-native English-speaking teachers' challenges in Taiwanese elementary schools. *System*, 38(4), 443-455. Retrieved from <https://doi.org/10.1016/j.system.2009.12.004>
- Chien, C.-W. (2022). Facilitating Taiwanese elementary school English student teachers' learning through text-based protocols. *Australian Journal of Teacher Education*, 47(10), Article 1. Retrieved from <https://doi.org/10.14221/ajte.2022v47n10.1>
- Cvitkovic, R., & Kita, Y. (2025). An overview of team teaching in Japanese

English education: A literature review of successes, challenges, and future directions. *Teikyo Journal of Language Studies*, 16, 119-150. Retrieved from <https://teikyo-u.repo.nii.ac.jp/record/2068063/files/gaikokubunka16-05.pdf>

■ Graham, K. M., & Yeh, Y.-F. (2022). Teachers' implementation of bilingual education in Taiwan: Challenges and arrangements. *Asia Pacific Education Review*. *Advance online publication*. Retrieved from <https://doi.org/10.1007/s12564-022-09791-4>

■ Kim, S.-Y., & Moodie, I. (2023). The models of co-teaching and a spectrum for assessing collaboration: Examining English language co-teaching practices in South Korea. *Language Teaching Research*. *Advance online publication*. Retrieved from <https://doi.org/10.1177/13621688231218816>

■ Lee, J. M. (2021). Claims of entitlements in elementary EFL co-teaching: A discourse-analytic study. *English Teaching*, 76(3), 63-84. Retrieved from <https://doi.org/10.15858/engtea.76.3.202109.63>

■ Lin, C.-C., & Wu, M.-H. (2025). Navigating Taiwan's Bilingual 2030 policy: Unveiling teachers' glocalization efforts and agency. *Journal of Multilingual Education Research*, 13, 83-107. Retrieved from <https://doi.org/10.5422/jmer.2024-2025.v13.83-107>

■ Luo, W.-H. (2007). A study of native English-speaking teacher programs in elementary schools in Taiwan. *Asia Pacific Education Review*, 8, 311-320. Retrieved from <https://doi.org/10.1007/BF03029265>

■ Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., ... Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, 372, n71. Retrieved from <https://doi.org/10.1136/bmj.n71>

■ Park, J. E. (2014). English co-teaching and teacher collaboration: A micro-interactive perspective. *System*, 44, 34-47. Retrieved from <https://doi.org/10.1016/j.system.2014.02.003>

■ Shin, H.-W., Lee, J., Brawn, J., & Do, J. (2021). EPIK teachers' beliefs about language learning and teaching processes. *English Teaching*, 76(2), 81-105. Retrieved

from <https://doi.org/10.15858/engtea.76.2.202106.81>

- Smith, E. (2021). Effectiveness of English teaching with JET Programme: Team teaching perceptions through team interviews [Master's thesis, Stockholm University]. DiVA Portal. Retrieved from <https://www.diva-portal.org/smash/get/diva2:1574558/FULLTEXT01.pdf>
- 教育部。(2018)。2030 雙語國家政策宣示 [Policy statement on the 2030 Bilingual Nation]。國家發展委員會。取自 <https://bilingual.ndc.gov.tw/>
- 國家發展委員會。(2022)。2030 雙語政策（政策總覽） [Bilingual 2030 policy (overview)]。Retrieved from <https://bilingual.ndc.gov.tw/>
- Critical Appraisal Skills Programme. (2018). CASP checklists. 取自 <https://casp-uk.net/casp-tools-checklists/>
- Hong, Q. N., Pluye, P., Fàbregues, S., Bartlett, G., Boardman, F., Cargo, M., Dagenais, P., Gagnon, M.-P., Griffiths, F., Nicolau, B., O' Cathain, A., Rousseau, M.-C., Vedel, I., & Zarowsky, C. (2018). Mixed Methods Appraisal Tool (MMAT) version 2018: User guide. McGill University, Department of Family Medicine. Retrieved from <http://mixedmethodsappraisaltoolpublic.pbworks.com/>
- Fulbright Taiwan (Foundation for Scholarly Exchange). (2024-2025). Program book (2024-2025) [PDF]. Retrieved from <https://www.fulbright.org.tw/wp-content/uploads/2024/11/2024-2025-program-book.pdf>

