

Appendix 1 Study Context and Community Background

This study was conducted in a small private school located in Antipolo City, Philippines, serving a predominantly low SES community. According to the Philippine Statistics Authority (PSA), the poverty incidence in Rizal Province, where Antipolo is situated, was approximately 7.2% as of the most recent survey, with localized disparities in urban and rural areas. However, the community served by the school experiences higher poverty levels, with over 30% of families reporting incomes below the official poverty threshold of PHP 12,030 per month for a family of five.

The school itself reflects these economic constraints. Tuition fees are kept deliberately low, averaging PHP 5,000–7,000 per year, to accommodate families with limited financial capacity. Despite this, many students benefit from financial aid or installment payment plans, indicating the economic challenges faced by the community. Furthermore, school facilities are minimal, with limited access to technology and laboratory resources, which aligns with the socio-economic profile of the population served.

These statistics and contextual details substantiate the characterization of the community as low SES and provide a clear basis for examining how socio-economic factors intersect with cultural and linguistic influences on the implementation of the LS framework.

Appendix 2 Details of Data Collection

Focus Group Discussions (FGDs)

FGDs were a key data collection method in this study, designed to explore teachers' experiences, expectations, and the cultural and linguistic insights that influenced their engagement with the LS framework. A total of five FGDs were conducted: one at the beginning of the study to establish baseline perceptions and four after each LS cycle to gather reflections on the process. Each session lasted approximately 60–90 minutes and was held in a private, quiet space at the school to ensure confidentiality and minimize distractions.

The questions for the FGDs were developed based on a combination of the study's research objectives, existing literature on LS and professional development, and cultural and linguistic considerations specific to the Philippine context. The goal was to elicit meaningful and focused discussions that would shed light on the following themes: teachers' perceptions of LS and their expectations before implementation, the influence of cultural values (e.g., *pakikisama* and *hiya*) on collaboration and feedback, the role of bilingualism in shaping communication and reflections during LS cycles; challenges and benefits experienced during the implementation of LS, suggestions for adapting LS to better align with the local context.

To ensure the questions were clear, relevant, and culturally appropriate, they were pre-tested with a small group of educators not involved in the study. Feedback from this pre-test was used to refine the questions further. The following is the final set of FGD questions used during the study:

Baseline FGD (Pre-implementation):

1. What are your initial thoughts on Lesson Study? Have you heard of or participated in similar professional development models?
2. What do you expect to gain from participating in this process?
3. How do you perceive collaboration and feedback in your current teaching practices?

FGDs After Each LS Cycle:

1. What were your experiences during this cycle of LS? Were there any challenges you encountered?
2. How did cultural values (e.g., *pakikisama*, *hiya*) influence your participation in planning and reflection sessions?
3. How did you navigate language use (Filipino and English) during discussions? Did it impact the depth of your reflections?
4. What strategies or practices during this cycle worked well? What could be improved for the next cycle?

Final FGD (Post-implementation):

1. Looking back at all the LS cycles, how would you describe the overall impact on your teaching practices?
2. How do you feel the cultural and linguistic aspects of your work environment influenced the LS process?
3. What recommendations would you provide for improving the implementation of LS in similar contexts?

Audio recordings and detailed notes were used to document the discussions, and participant consent for recording was obtained before each session. The structured but open-ended nature of the questions allowed for deep exploration of the study themes while encouraging participants to share their thoughts freely. These discussions provided rich qualitative data, forming a critical component of the study's findings.

Classroom Observations

Classroom observations were conducted to capture real-time teacher-student and teacher-peer interactions during Lesson Study (LS) cycles. A total of eight observations were carried out, with two observations per LS cycle. Each observation lasted approximately 45–60 minutes, corresponding to the length of the lessons being implemented. The observations were performed by researchers using a structured observational guide to ensure consistency and reliability in data collection. The following outlines the details of the observational guide:

1. *Teacher Practices and Pedagogical Strategies*
 - How does the teacher introduce and explain concepts?
 - What strategies are used to engage students (e.g., questioning, group activities)?
 - Are the lesson objectives clearly communicated and aligned with the activities?
2. *Student Engagement and Responses*
 - How actively are students participating in the lesson (e.g., asking questions, collaborating with peers)?
 - Do students appear to understand the concepts being taught (e.g., correct responses, active problem-solving)?
 - Are there any observable differences in engagement across different groups of students (e.g., gender, language proficiency)?
3. *Collaborative and Reflective Practices (Post-Observation Reflection)*
 - How do teachers interact with their peers during post-lesson reflections?
 - Are teachers able to provide and receive constructive feedback?
 - What cultural or linguistic factors (e.g., use of Filipino or English, influence of *pakikisama* or *hiya*) are evident during reflective discussions?

Researchers paid close attention to the following areas during observation: (1) the clarity of the teacher's explanations and the alignment between lesson plans and classroom implementation, (2) specific examples and analogies used by the teacher to explain concepts, noting any reliance on simplified examples, (3) instances of bilingual communication, including code-switching between Filipino and English, and its effect on both teaching and student understanding, (4) student reactions, including their engagement levels, verbal responses, and non-verbal cues like expressions or gestures, and (5) interaction dynamics during collaborative activities, such as group work or peer feedback sessions.

All observations were audio- and video-recorded to capture both verbal and non-verbal interactions comprehensively. Detailed field notes supplemented the recordings, focusing on elements such as teacher and student behavior, the physical classroom environment, and any notable disruptions or contextual factors. The observational guide was applied systematically to ensure consistency across all sessions.

These structured observations provided valuable insights into how teachers implemented LS, how students engaged with the lessons, and how cultural and linguistic factors influenced

classroom and post-lesson dynamics. The guide ensured that the data collected was both detailed and relevant, contributing significantly to the analysis and findings of the study.

Individual Interviews

Semi-structured interviews were conducted with each teacher at the end of every LS cycle to elicit personal reflections on their experiences and to delve deeper into themes that emerged during FGDs and observations. Each interview lasted 30–60 minutes and followed a guide designed to balance consistency with flexibility, allowing participants to elaborate on their perspectives. The interview questions focused on cultural and linguistic factors, collaborative dynamics, and the perceived impact of LS on their teaching practices. Researchers took measures to minimize bias by avoiding leading questions and maintaining a neutral tone. Additionally, cultural sensitivities, such as the influence of *hiya* (modesty) and *pakikisama* (harmony), were acknowledged, with participants reassured that their responses would remain confidential. All interviews were audio-recorded, and transcription was done verbatim to preserve the authenticity of responses.

Lesson Artifacts and Reflective Documents

Teachers' lesson plans, instructional materials, and student work samples were collected as part of the data set to provide additional insights into how the LS process influenced planning and instructional practices. These artifacts were analyzed in conjunction with other data sources to identify recurring themes related to professional growth and cultural considerations.

Appendix 3 Details of Data Analysis

Coding Approach and Framework

The data analysis utilized a combination of inductive and deductive coding, guided by the research questions and the theoretical framework of the study. Deductive codes were informed by established literature on LS and culturally responsive professional development, particularly the frameworks proposed by Lewis, Perry, and Hurd (2009) and Mukwambo et al. (2018). These codes included categories such as "collaboration," "reflective practice," and "feedback dynamics." Concurrently, inductive codes emerged during the analysis process, capturing themes unique to the Philippine context, such as "hesitancy in critique," "code-switching," and "SIR." The development and refinement of codes were systematically documented in a codebook, which provided detailed definitions, examples, and application rules to ensure consistency.

Data Management and Coding Tools

All data, including transcripts of focus group discussions FGDs and interviews, as well as detailed observation notes, were managed and analyzed using Microsoft Excel. Although not a specialized qualitative analysis tool, Excel was selected for its accessibility and flexibility in organizing and categorizing qualitative data. Verbatim transcripts were uploaded, and audio and video recordings were reviewed to annotate verbal and non-verbal cues, such as tone, pauses, and gestures, which were critical for understanding the depth and context of participant interactions. Non-verbal observations were linked to the transcripts to provide richer insights into the data.

Inductive Coding Process

The coding process involved three distinct stages: initial coding, axial coding, and selective coding. In the initial phase, transcripts and observation notes were reviewed line by line to identify preliminary codes. Deductive codes were applied to text segments aligning with theoretical constructs, while inductive codes emerged from patterns observed within the data. For instance, the code "hesitancy in critique" reflected the influence of cultural values, such as *hiya* and *pakikisama*, on teachers' willingness to engage in feedback. Verbal and non-verbal cues were coded separately, with annotations for behaviors such as gestures or expressions complementing verbal data.

During the axial coding phase, codes were grouped into broader categories based on observed relationships and patterns. For example, codes such as "*pakikisama*," "*hiya*," and "self-restraint" were grouped under the theme "Cultural Sensitivity in Feedback," which captured how cultural norms shaped teacher interactions. Similarly, "language preference," "code-switching," and "comfort with English" were consolidated under "Impact of Language on Communication." The refinement process ensured that emerging themes accurately reflected the data, with adjustments made iteratively to improve clarity and coherence.

In the final phase, selective coding focused on overarching themes central to the research questions, supported by representative codes. The frequency and distribution of these codes were analyzed across data sources, such as FGDs, interviews, and classroom observations, to identify commonalities and differences. For instance, the code "hesitancy in critique" appeared more frequently in FGDs (42 instances) compared to interviews (16 instances), highlighting group-level dynamics that influenced reflective practices.

Table 3.1 summarizes the key themes, their descriptions, and examples of codes, with excerpts from the data sources to illustrate their application. The table also specifies whether the data came from FGDs, interviews, or observations, providing transparency and clarity.

Table 3.1 Overview of the key themes

Theme	Description	Codes	Example Excerpts	Source
Cultural Sensitivity in Feedback	How cultural values, such as <i>pakikisama</i> and <i>hiya</i> , influence teachers' comfort and willingness to engage in critical discussions and provide feedback.	<i>Pakikisama, Hiya, Self-restraint, Avoiding conflict</i>	<i>"I didn't want to say too much about her lesson because it might sound too critical."</i>	FGDs, Interviews
Impact of Language on Communication	The role of bilingualism and language preference in shaping teachers' comfort and depth of engagement in LS discussions.	Language preference, Comfort with English, Code-switching	<i>"I felt I could express myself better in Filipino, especially during feedback."</i>	FGDs, Interviews
Professional Growth through Reflection	Teachers' reflections on how LS has contributed to their development in terms of pedagogical knowledge and collaborative learning.	Self-reflection, Pedagogical insights, Growth in teaching	<i>"I learned to think more about why I choose certain activities, not just what's easy."</i>	Interviews
Challenges with Content Knowledge	Teachers' difficulties in articulating theoretical principles and demonstrating deep content knowledge during lesson planning and reflection.	Content knowledge gaps, Simplified examples, Theoretical clarity	<i>"I used pizza as an example for fractions, but I realize now it's limiting; I'm not sure what else I could use, though."</i>	Observations
Supportive Peer Relationships	The role of mutual respect, trust, and collegial support among teachers, enabling a collaborative atmosphere for LS despite cultural constraints.	Trust-building, Collegiality, Mutual support	<i>"I felt more comfortable sharing my struggles with them because I knew they wouldn't judge me."</i>	FGDs, Interviews

Inter-Rater Reliability

To ensure reliability, three researchers (the two authors plus an independent coder) independently coded 20% of the data and compared results before coding the rest of the data. Discrepancies were discussed collaboratively, refining code definitions and applications. Inter-rater reliability was calculated using Cohen's kappa, with an average score of 0.82, indicating substantial agreement and ensuring consistency in the analysis (McHugh, 2012).

Code Frequencies and Distribution

The frequency and distribution of codes across FGDs, interviews, and observations were calculated to identify patterns. For example, "Cultural Sensitivity in Feedback" was observed most frequently in FGDs, reflecting group-level dynamics, while "Professional Growth through Reflection" appeared more prominently in individual interviews. These frequencies were also analyzed by teacher, discipline, and LS cycle to explore variations across different contexts, offering a nuanced understanding of how cultural and linguistic factors influenced LS implementation.