

PROCEEDING OF INTERNATIONAL SEMINAR GLOBAL ISSUES IN SOCIAL, SCIENCE, LAW, HEALTH AND BUSINESS ENVIRONMENT

Is Evaluation Through Lesson Study Feasible in Indonesia? A Reflection on Theory-Based Evaluation of Lesson Study Professional Development: Challenges, Opportunities, and Lessons Learned by D. Jake Follmer et al.

Niko Sudibjo¹, HG Retno Harsanti^{1,2}

¹Universitas Pelita Harapan

²National Taiwan Normal University

Corresponding: niko.sudibjo@uph.edu

Abstract

This review paper examines the feasibility of evaluating Lesson Study in Indonesia through the lens of theory-based evaluation, reflecting on the framework proposed by Follmer et al. (2024). The article argues that complex instructional interventions such as Lesson Study require theoretically grounded, multi-level, and temporally sensitive evaluation models capable of capturing recursive and socially regulated learning processes. While this framework demonstrates strong conceptual rigor and empirical promise, its application in Indonesia requires contextual adaptation. Indonesia's collectivist culture may support collaborative inquiry; however, hierarchical norms, high uncertainty avoidance, and compliance-oriented supervision systems may constrain open critique, iterative reflection, and theory-driven inquiry. Intensive classroom observation risks Hawthorne effects and increased cognitive load for students, potentially distorting authentic learning evidence. Institutional workload pressures may also lead to superficial collaboration rather than genuine professional learning. The paper concludes that Lesson Study-based evaluation is feasible in Indonesia only when supported by strong instructional leadership, psychological safety, and contextualized implementation. Rather than a standardized tool, it should be strategically adapted as a culturally responsive professional learning practice.

Keywords: Theory-Based Evaluation, Lesson Study, Professional Development, Socially Regulated Learning, Indonesian Education Context

Introduction

Follmer et al. (2024) address a persistent challenge in educational evaluation: the limited and inconsistent use of theory to guide the evaluation of complex instructional programs. The authors argue that while evaluation findings are central to shaping educational practice, informing accountability, and engaging stakeholders, existing evaluation models often struggle to capture the multidimensional and process-oriented nature of instructional interventions. In response, the article advances the case for theory-based evaluation as a necessary framework for assessing instructional programs that are integrated, iterative, and contextually embedded. The authors contend that effective instructional evaluation must be theoretically grounded, capable of examining outcomes across multiple levels, and sensitive to temporal dynamics in both processes and effects.

The Review of the lesson study

Situating Lesson Study in Theory-Based Evaluation

PROCEEDING OF INTERNATIONAL SEMINAR GLOBAL ISSUES IN SOCIAL, SCIENCE, LAW, HEALTH AND BUSINESS ENVIRONMENT

The article situates lesson study within the broader tradition of theory-based evaluation, which emphasizes the explicit articulation of causal models linking program inputs and activities to intended or observed outcomes (Rogers, 2008). Follmer et al. (2024) highlight that theory-based evaluation requires not only the selection of appropriate theoretical frameworks but also careful attention to the quality and coherence of program theories used to explain how and why interventions work. The authors acknowledge longstanding challenges in this approach, including difficulties in theory selection, the construction of accurate causal models, intensive data demands, and the time required to test and refine theoretical assumptions. Despite these challenges, they argue that theory-based evaluation is particularly well suited to lesson study, as it enables evaluators to account for program complexity, contextual variation, and the interplay between instructional processes and outcomes.

Evaluation of Lesson Study Current Research and Gaps

The authors review existing research on lesson study evaluation and note that, although substantial evidence supports its effectiveness, evaluation practices remain uneven and methodologically limited. Prior studies have primarily focused on proximal outcomes, such as teachers' beliefs, attitudes, professional learning, instructional practices, and, to a lesser extent, student learning and school-level professional development structures. Follmer et al. (2024) argue that these approaches often fail to capture the full range of lesson study's impacts, particularly its systemic and longitudinal effects. They call for more rigorous, theory-driven, and system-oriented evaluation designs capable of addressing lesson study's recursive and collaborative nature.

Theoretical Underpinnings of Our Approach to Lesson Study Evaluation

The evaluation framework proposed in the article is grounded in both cognitive and situated learning theories. Lesson study is conceptualized as a process that supports knowledge restructuring while simultaneously fostering professional identity development within a community of practice. The authors observe that lesson study research has largely converged on a shared implementation theory encompassing lesson investigation, collaborative planning, enactment, observation, and reflective analysis. Building on this foundation, the authors articulate a program theory informed by self-regulated learning theory (Zimmerman) and socially shared regulation of learning. This theoretical integration positions teachers as both learners and instructors engaged in cyclical processes of goal setting, performance monitoring, collective reflection, and continuous improvement within professional learning communities.

Implementing a Socially Regulated Theory-Based Evaluation

Empirically, the study examines a lesson study initiative implemented within a public university's accredited teacher preparation program in the Mid-Atlantic United States. The program aims to prepare and retain highly qualified teachers for high-need schools by integrating lesson study into preservice preparation and ongoing professional development for in-service teacher mentors. Lesson study activities are conducted at school placements involving middle and high school STEM educators, followed by structured reflection and debriefing sessions at the university. This design allows the evaluation to capture both individual-level learning processes and broader professional development outcomes.

PROCEEDING OF INTERNATIONAL SEMINAR GLOBAL ISSUES IN SOCIAL, SCIENCE, LAW, HEALTH AND BUSINESS ENVIRONMENT

Causal Pathways and Program Theory

A central contribution of the article lies in its delineation of complex causal strands linking lesson study components, such as content knowledge development, pedagogical planning, and student assessment, to outcomes including enhanced teaching efficacy and instructional quality. These causal pathways account for both direct and indirect influences, incorporating participants' prior experiences, recursive practice cycles, and contextual moderators such as program duration, facilitation quality, teacher beliefs, and group dynamics. By integrating cognitive, affective, and contextual dimensions aligned with self- and socially regulated learning, the authors propose a comprehensive program theory that moves beyond traditional, linear evaluation models. Notably, the inclusion of affective variables such as motivation and self-efficacy deepens the analysis of teacher learning processes within lesson study contexts.

Outcomes, Measures, and Analytical Strategies

To align outcomes with the theoretical framework, the authors employ a mixed-methods approach drawing on multiple sources of evidence. These include qualitative and quantitative participant reports, lesson study artifacts, assessments of teaching knowledge and beliefs, coded debriefing sessions, analyses of pedagogical reasoning and reflection, and documented refinements in lesson planning. Analytical strategies range from descriptive and within-person change analyses to deductive and inductive qualitative methods. Outcomes are examined across different phases of the lesson study cycle and at multiple levels, including individual teachers, professional learning communities, and school systems. This alignment enables the evaluation to trace how specific processes contribute to observed outcomes over time.

Key Findings

The findings indicate that lesson study positively influences teachers' self-efficacy, instructional effectiveness, and engagement in reflective practice, with downstream benefits for student engagement. Structured debriefing sessions, particularly those conducted in small groups and whole-group formats, are shown to enhance teachers' expectations for student learning and deepen content understanding. The transition to online debriefing during the COVID-19 pandemic further revealed the evaluative potential of digital platforms, which facilitated synchronous discussion analysis, video-based data collection, and richer qualitative insights.

Implications and Lessons Learned

Follmer et al. (2024) conclude that lesson study generates immediate and meaningful changes in teachers' beliefs and instructional practices, with implications for both classroom-level teaching and school-level decision-making. They emphasize that the iterative and inquiry-driven nature of lesson study necessitates theory-based evaluation approaches capable of capturing recursive causality and non-linear effects. The authors argue for the integration of quantitative and qualitative data to move beyond static outcome measures and better represent dynamic group interactions. Ultimately, the socially regulated evaluation framework advanced in this study provides a robust interpretive lens for understanding how lesson study functions as a complex professional learning intervention, while also offering guidance for adapting theory-based evaluation to similarly multifaceted instructional programs.

PROCEEDING OF INTERNATIONAL SEMINAR GLOBAL ISSUES IN SOCIAL, SCIENCE, LAW, HEALTH AND BUSINESS ENVIRONMENT

Reflection on the Feasibility of Evaluation through Lesson Study in the Indonesian Context

While the theory-based evaluation of lesson study proposed by Follmer et al. (2024) offers a conceptually robust framework for examining instructional improvement, its feasibility in the Indonesian educational context requires careful theoretical and contextual consideration. Indonesia's education system is shaped by distinct cultural, institutional, and pedagogical characteristics that may simultaneously support and constrain the implementation of lesson study as an evaluative approach.

From a cultural standpoint, Indonesia is commonly characterized, based on Hofstede's cultural dimensions as a society with strong collectivist orientations, relatively low competitiveness, and a high tendency toward uncertainty avoidance (Hofstede et al., 2010). The collectivist orientation potentially aligns well with the collaborative foundations of lesson study, which emphasize shared inquiry, collective responsibility, and professional learning communities. However, high uncertainty avoidance may present a significant challenge. Theory-based evaluation and lesson study both rely on iterative inquiry, openness to critique, and tolerance for ambiguity, conditions that may be perceived as destabilizing in educational cultures that value procedural clarity, predictability, and hierarchical accountability (Hall & Hord, 2015). As a result, lesson study may be accepted normatively while resisted substantively, particularly when its outcomes are not immediately tied to formal evaluation or administrative recognition.

A central practical concern lies in the intensive classroom observation inherent in lesson study. In Indonesia, instructional observation involving multiple observers within a single classroom is rare and may generate discomfort for both teachers and students. From a behavioral research perspective, such observation risks producing a Hawthorne effect, whereby individuals alter their behavior simply because they are aware of being observed (Mayo, 2010). For students, this may result in heightened compliance, anxiety, or performative engagement that does not accurately represent authentic learning processes. Consequently, observational data collected under such conditions may overestimate instructional effectiveness or mask underlying learning difficulties.

These concerns are further amplified when viewed through the lens of students' psychological comfort and cognitive processing. Cognitive load theory Sweller (2011) suggests that extraneous stimuli, such as the presence of unfamiliar adults or heightened evaluative pressure, can increase extraneous cognitive load, thereby reducing students' capacity to engage meaningfully with instructional content. Given that Indonesian students, at the system level, are still striving to improve foundational competencies in literacy and numeracy, they may be particularly sensitive to such disruptions. In this context, lesson study-based evaluation, if not carefully adapted, may inadvertently interfere with learning rather than support it.

Institutionally, evaluation is not absent from Indonesian education, as supervision mechanisms are formally embedded within the system. However, existing supervisory practices tend to prioritize administrative compliance, documentation, and procedural accountability rather than deep instructional analysis (OECD, 2023). Lesson study represents a fundamentally different evaluative logic, one that centers on instructional reasoning, student thinking, and iterative pedagogical refinement. From the perspective of evaluation theory, this aligns with formative and developmental evaluation approaches (Patton, 2011), which emphasize learning-oriented feedback rather than summative judgment. Nevertheless, the collaborative demands of

PROCEEDING OF INTERNATIONAL SEMINAR GLOBAL ISSUES IN SOCIAL, SCIENCE, LAW, HEALTH AND BUSINESS ENVIRONMENT

lesson study, from joint planning to collective reflection, require significant time, cognitive investment, and relational trust.

In practice, heavy administrative workloads and fragmented professional development structures may lead to what Hargreaves (2000) terms contrived collegiality, where collaboration exists formally but lacks depth and authenticity. Teachers may engage in lesson study activities superficially, driven by compliance rather than professional inquiry. This tension underscores the importance of organizational conditions in enabling socially regulated learning (Hadwin et al., 2017), where collective goal-setting, shared monitoring, and mutual regulation are genuinely enacted rather than symbolically performed.

Theoretically, these challenges highlight the central role of instructional leadership and distributed leadership frameworks in mediating the feasibility of lesson study (Robinson, 2011; Spillane, 2006). School leaders are pivotal in establishing psychological safety (Edmondson, 1999), reframing observation as a developmental rather than evaluative act, and aligning lesson study with teachers' existing responsibilities. When leaders actively cultivate trust, normalize uncertainty, and protect instructional time, lesson study can function as a powerful mechanism for professional learning and theory-based evaluation rather than an additional administrative burden.

In conclusion, lesson study constitutes a valuable and theoretically grounded evaluative practice with significant potential to enhance instructional quality in Indonesia. However, its feasibility is contingent upon careful cultural, psychological, and institutional adaptation. Rather than being adopted as a standardized evaluative tool, lesson study should be approached as a good practice worthy of critical learning and contextual translation. In particular, school leaders must develop a deep theoretical and practical understanding of lesson study to guide its implementation strategically, mitigating unintended negative effects while maximizing its capacity to support meaningful, sustained improvements in teaching and learning.

References

- Edmondson, A. (1999). Psychological Safety and Learning Behavior in Work Teams. *Administrative Science Quarterly*, 44(2), 350–383. <https://doi.org/10.2307/2666999>
- Follmer, D. J., Groth, R., Bergner, J., & Weaver, S. (2024). Theory-based Evaluation of Lesson Study Professional Development: Challenges, Opportunities, and Lessons Learned. *American Journal of Evaluation*, 45(2), 292–312. <https://doi.org/10.1177/10982140231184899>
- Hadwin, A., Järvelä, S., & Miller, M. (2017). Self-Regulation, Co-Regulation, and Shared Regulation in Collaborative Learning Environments. In *Handbook of Self-Regulation of Learning and Performance* (2nd ed.). Routledge.
- Hall, G. E., & Hord, S. M. (2015). *Implementing Change: Patterns, Principles, and Potholes*. Pearson.
- Hargreaves, A. (2000). Four Ages of Professionalism and Professional Learning. *Teachers and Teaching*, 6(2), 151–182. <https://doi.org/10.1080/713698714>
- Hofstede, G., Hofstede, G. J., & Minkov, M. (2010). *Cultures and Organizations: Software of the Mind, Third Edition*. McGraw Hill Professional.

PROCEEDING OF INTERNATIONAL SEMINAR GLOBAL ISSUES IN SOCIAL, SCIENCE, LAW, HEALTH AND BUSINESS ENVIRONMENT

Mayo, E. (2010). *The Human Problems of an Industrial Civilization*. Routledge.

OECD. (2023). *PISA 2022 Results Factsheets Indonesia*. <https://www.oecd.org/publication/pisa-2022-results/webbooks/dynamic/pisa-country-notes/c2e1ae0e/pdf/indonesia.pdf>

Patton, M. Q. (2011). *Developmental Evaluation: Applying Complexity Concepts to Enhance Innovation and Use*. The Guilford Press.

Robinson, V. (2011). *Student-Centered Leadership*. Jossey-Bass.

Rogers, P. J. (2008). Using Programme Theory to Evaluate Complicated and Complex Aspects of Interventions. *Evaluation*, 14(1), 29–48. <https://doi.org/10.1177/1356389007084674>

Spillane, J. P. (2006). *Distributed Leadership*. Wiley.

Sweller. (2011). *Cognitive Load Theory*. Springer.