



DOI: <https://doi.org/10.38035/jmpis.v7i5>  
<https://creativecommons.org/licenses/by/4.0/>

## Evaluasi Program Supervisi Akademik Berbasis *Lesson Study* di SD Negeri 1 Marga Kabupaten Tabanan

I Nyoman Agus Triwijaya<sup>1</sup>, Kadek Rihendra Dantes<sup>2</sup>, I Putu Wisna Ariawan<sup>3</sup>

<sup>1</sup>Universitas Pendidikan Ganesha, Singaraja, Indonesia, [agus.triwijaya@student.undiksha.ac.id](mailto:agus.triwijaya@student.undiksha.ac.id)

<sup>2</sup>Universitas Pendidikan Ganesha, Singaraja, Indonesia, [rihendra-dantes@undiksha.ac.id](mailto:rihendra-dantes@undiksha.ac.id)

<sup>3</sup>Universitas Pendidikan Ganesha, Singaraja, Indonesia, [wisna.ariawan@undiksha.ac.id](mailto:wisna.ariawan@undiksha.ac.id)

Corresponding Author: [agus.triwijaya@student.undiksha.ac.id](mailto:agus.triwijaya@student.undiksha.ac.id)<sup>1</sup>

**Abstract:** *This study evaluates the Lesson Study-based academic supervision program at SD Negeri 1 Marga, Tabanan Regency, a school that has run the program for three years yet still records a moderate score on the Learning Quality indicator despite gains in teacher reflection and instructional leadership. The study assesses program effectiveness through the Context, Input, Process, Product model, identifies factors explaining the gap between process strengthening and learning-quality stagnation, and formulates strategic recommendations for supervisors, the principal, teachers, and the local education office. Using a descriptive quantitative evaluation research design, data were collected from all ten stakeholders at the school (eight teachers, one principal, one supervisor) through a Likert-scale questionnaire, interviews, classroom observation, and document study, then analyzed using score tabulation, Z-score and T-score transformation, and Glickman quadrant mapping. Results show the program is rated Good to Highly Effective at the aggregate school level, but effectiveness is unevenly distributed among individuals: the Input component is weakest, and three of eight teachers fall into the Less Effective category solely due to below-cutoff Product scores, including one teacher who paradoxically shows strong process engagement. The study concludes that process strengthening does not automatically translate into consistent classroom-level improvement, requiring periodic input evaluation and individualized post-reflection mentoring.*

**Keywords:** *Academic Supervision, Lesson Study, CIPP Evaluation Model, Instructional Leadership, Learning Quality*

**Abstrak:** Penelitian ini mengevaluasi Program Supervisi Akademik Berbasis *Lesson Study* di SD Negeri 1 Marga, Kabupaten Tabanan, sekolah yang telah menjalankan program tersebut selama tiga tahun, namun masih mencatat capaian sedang pada indikator Kualitas Pembelajaran meskipun mengalami penguatan pada indikator refleksi guru dan kepemimpinan instruksional. Penelitian bertujuan mengevaluasi efektivitas program melalui model *Context, Input, Process, Product*, mengidentifikasi faktor yang menjelaskan kesenjangan antara penguatan proses dan stagnasi kualitas pembelajaran, serta merumuskan rekomendasi strategis bagi pengawas, kepala sekolah, guru, dan dinas pendidikan setempat. Menggunakan desain

evaluasi dengan pendekatan kuantitatif deskriptif, data dikumpulkan dari seluruh sepuluh pemangku kepentingan di sekolah tersebut (delapan guru, satu kepala sekolah, satu pengawas) melalui kuesioner berskala Likert, wawancara semi-terstruktur, observasi pembelajaran, dan studi dokumentasi, kemudian dianalisis melalui tabulasi skor mentah, transformasi Z-skor dan T-skor, serta pemetaan kuadran Glickman. Hasil penelitian menunjukkan program tergolong Baik hingga Sangat Efektif pada level agregat sekolah, tetapi efektivitas tersebut tidak terdistribusi merata pada level individu: komponen Masukan menjadi komponen paling lemah, dan tiga dari delapan guru berada pada kategori Kurang Efektif semata-mata karena skor Produk di bawah titik potong, termasuk satu guru yang justru menunjukkan keterlibatan proses yang kuat. Penelitian ini menyimpulkan bahwa penguatan proses tidak otomatis terkonversi menjadi perbaikan konsisten di level kelas, sehingga diperlukan evaluasi masukan secara berkala dan pendampingan reflektif individual, bukan sekadar perpanjangan program secara seragam.

**Kata Kunci:** Supervisi Akademik, *Lesson Study*, Model Evaluasi CIPP, Kepemimpinan Instruksional, Kualitas Pembelajaran

## PENDAHULUAN

Kualitas pembelajaran di sekolah dasar sangat bergantung pada pembinaan profesional guru. Melalui Standar Pengawas Sekolah/Madrasah, supervisi akademik berfungsi sebagai instrumen formal untuk memastikan praktik mengajar guru konsisten dengan tujuan pembelajaran. Sejak 2023, SD Negeri 1 Marga menjadi salah satu sekolah inti di Gugus 3 Kecamatan Marga yang menerapkan supervisi akademik berbasis *Lesson Study* melalui siklus *plan-do-see* yang disinergikan dengan Penilaian Kinerja Guru, sehingga secara desain sekolah ini diharapkan menjadi lokomotif peningkatan mutu pembelajaran di gugus. Namun, data Rapor Pendidikan Kemendikdasmen 2025 (Permendikdasmen No. 1, 2025) justru memunculkan problem empiris yang jelas: indikator Refleksi dan Perbaikan Guru (D.2) di gugus naik 6,52 poin dan Kepemimpinan Instruksional (D.3) naik 8,70 poin, sementara Kualitas Pembelajaran (D.1) sebagai indikator hasil hanya bergerak 0,22 poin; SDN 1 Marga dan SDN 1 Kuwum bahkan masih berkategori sedang pada D.1. Selain itu, kesenjangan lebih dari 27 poin pada sub-indikator Dukungan Refleksi Guru (D.3.3) antarsekolah menunjukkan bahwa penguatan proses (refleksi, pendampingan, kepemimpinan) terdistribusi secara tidak merata dan belum terkonversi menjadi perbaikan kualitas pembelajaran yang terukur di SDN 1 Marga.

Kondisi tersebut mengarah pada problem statement yang spesifik: setelah tiga tahun implementasi supervisi akademik berbasis *Lesson Study*, SDN 1 Marga mengalami peningkatan signifikan pada dimensi proses (D.2 dan D.3), tetapi kualitas pembelajaran di kelas (D.1) tetap stagnan di kategori sedang. Artinya, terdapat kesenjangan proses–produk di tingkat sekolah, yakni antara kapasitas reflektif guru dan dukungan kepemimpinan di satu sisi dengan praktik pembelajaran yang berdampak di sisi lain. Dalam perspektif evaluasi program pendidikan, kesenjangan ini bukan sekadar anomali angka, melainkan indikator bahwa desain dan pelaksanaan program belum berhasil menjadikan *Lesson Study* dan supervisi akademik sebagai mekanisme yang efektif untuk mentransformasikan peningkatan kapasitas guru menjadi perubahan praktik mengajar yang konsisten di SDN 1 Marga.

Secara teoritis, *Lesson Study* dipahami sebagai model pengembangan profesional berbasis kolaborasi yang berakar pada praktik *jugyo kenkyu* di Jepang, di mana guru bersama-sama merancang, melaksanakan, dan merefleksikan satu pembelajaran melalui observasi terbuka dan diskusi sistematis. Lewis menekankan bahwa siklus *plan-do-see* berfungsi bukan hanya sebagai perangkat teknis, melainkan sebagai cara membangun *professional learning*

*community* (PLC) yang belajar dari praktik kelas secara berulang. DuFour menegaskan bahwa PLC yang efektif menuntut prasyarat struktural (waktu kolaborasi yang terjadwal, dukungan organisasi) dan prasyarat kultural (kepercayaan, norma refleksi, dan tanggung jawab kolektif terhadap hasil belajar siswa), sehingga keberhasilan PLC tidak dapat direduksi menjadi frekuensi pertemuan atau keberadaan tim semata.

Dalam kerangka kepemimpinan instruksional Hallinger dan Murphy, kepala sekolah diposisikan sebagai pemimpin pembelajaran yang bertugas merumuskan tujuan akademik, mengkoordinasi kurikulum dan program pengembangan guru, memonitor proses pembelajaran, serta menyediakan dukungan profesional yang berkelanjutan. Penelitian Wahyuni et al. (2026) dan Egiska & Sucitra (2023) di konteks Indonesia menunjukkan bahwa kepemimpinan instruksional, termasuk ketika didistribusikan kepada pengawas dan guru senior sebagai *leadership for learning*, berhubungan positif dengan pengembangan profesional guru apabila dioperasionalkan melalui supervisi akademik dan pendampingan yang sistematis. Sementara itu, Ngatini et al. (2025) membuktikan bahwa kepemimpinan instruksional dan supervisi akademik secara simultan menjelaskan sebagian besar variasi kualitas pembelajaran di SMP, sehingga secara teoritis kedua konstruk ini diperkirakan mempengaruhi mutu pembelajaran melalui pengaruhnya terhadap kapasitas guru dan praktik mengajar di kelas.

Di sisi lain, kajian sistematis Indrawan et al., (2024) menunjukkan bahwa keberlanjutan dampak Lesson Study bersifat kontingen terhadap budaya kolegal dan dukungan kepemimpinan; siklus plan-do-see tidak otomatis menghasilkan perubahan praktik mengajar apabila PLC yang terbentuk tidak cukup kuat atau bila kepemimpinan instruksional lebih bersifat administratif daripada pedagogis. Ketegangan antara klaim keberhasilan Lesson Study yang bersifat generik (Astuti et al., 2026; Lewis, 2016; Salasiah et al., 2023; Sundari, 2023) dan temuan Indrawan et al., (2024) yang mensyaratkan kondisi kontekstual tertentu menyediakan lensa teoritis untuk membaca fenomena di SDN 1 Marga: peningkatan refleksi guru dan indikator kepemimpinan instruksional belum menjamin perbaikan kualitas pembelajaran apabila Lesson Study dan supervisi akademik dijalankan secara seremonial, atau bila hasil refleksi guru tidak terintegrasi secara konsisten ke dalam perencanaan dan eksekusi pembelajaran.

Jika ditelusuri, sebagian besar penelitian terdahulu menguji supervisi akademik, Lesson Study, kepemimpinan instruksional, dan PLC secara terpisah. Setyawati & Jaelani (2026) Menggunakan model CIPP untuk mengevaluasi supervisi akademik, tetapi belum mengintegrasikan Lesson Study sebagai strategi operasional dan belum memposisikan PLC sebagai variabel penjelas. Astuti et al., (2026) mengkaji peningkatan kompetensi guru berbasis Lesson Study secara deskriptif kualitatif tanpa evaluasi program yang formal, sehingga hubungan antara desain program, proses pelaksanaan, dan hasil pembelajaran tidak terbaca secara sistematis. Ichwan & Sukandar (2026) menunjukkan dampak positif supervisi kolaboratif reflektif terhadap kualitas pembelajaran, namun dimensi kepemimpinan instruksional tidak dibahas secara eksplisit sebagai faktor penentu keberhasilan. Ulum (2025) serta (Azizah & Nadia, 2024) memang mengevaluasi Lesson Study di sekolah dasar dan menemukan penguatan budaya reflektif dan kompetensi pedagogik guru, tetapi keduanya tidak menelusuri mengapa peningkatan kapasitas tersebut belum tentu berujung pada perubahan praktik mengajar yang terukur di kelas.

Di sisi kerangka evaluasi, literatur menunjukkan bahwa model CIPP yang dikembangkan Stufflebeam & Zhang (2017) merupakan salah satu model yang paling luas digunakan dalam evaluasi program pendidikan karena memisahkan empat dimensi utama: *Context*, *Input*, *Process*, dan *Product*. Berbagai tinjauan sistematis menggarisbawahi bahwa CIPP memberikan kerangka komprehensif untuk mengevaluasi program sekolah, tetapi masih terdapat celah berupa kurangnya integrasi dengan teori pengembangan profesional dan kepemimpinan instruksional, serta minimnya studi yang menelusuri hubungan sebab-akibat

antarkomponen secara mendalam di satu konteks sekolah. Dengan demikian, terdapat dua lapis *research gap*: pertama, belum ada kajian yang menyatukan supervisi akademik, *Lesson Study*, kepemimpinan instruksional, dan PLC dalam satu kerangka evaluasi berbasis CIPP; kedua, belum ada studi yang memfokuskan evaluasi semacam ini pada sekolah dasar dengan data empiris yang secara jelas menunjukkan kesenjangan proses produk seperti yang tampak di SDN 1 Marga.

Penelitian ini secara spesifik memposisikan SD Negeri 1 Marga sebagai kasus kunci untuk membaca bagaimana integrasi supervisi akademik, *Lesson Study*, kepemimpinan instruksional, dan PLC beroperasi di lapangan ketika dihadapkan pada data Rapor Pendidikan yang menampilkan peningkatan indikator proses (D.2 dan D.3) tetapi stagnasi indikator produk (D.1). Model evaluasi CIPP (*Context, Input, Process, Product*) digunakan untuk mengurai sumber kesenjangan tersebut: Context untuk menilai kesesuaian tujuan program dengan kebutuhan nyata guru dan sekolah; Input untuk menelaah desain *Lesson Study* dan supervisi akademik berdasarkan kerangka pengembangan profesional *Desimone (content focus, active learning, coherence, duration, collective participation)*; Process untuk menilai kualitas pelaksanaan siklus *plan-do-see* dan praktik supervisi akademik; serta Product untuk memeriksa kualitas pembelajaran di kelas sebagaimana terekam dalam indikator D.1.

Dengan kerangka itu, penelitian ini bertujuan: (1) mengevaluasi efektivitas Program Supervisi Akademik Berbasis *Lesson Study* di SD Negeri 1 Marga melalui model CIPP; (2) mengidentifikasi faktor kontekstual, kelemahan desain, dan dinamika pelaksanaan yang menjelaskan kesenjangan antara penguatan proses dan stagnasi kualitas pembelajaran; dan (3) merumuskan rekomendasi strategis bagi pengawas, kepala sekolah, guru, dan Dinas Pendidikan Kabupaten Tabanan terkait keberlanjutan, perbaikan, atau restrukturisasi program. Kontribusi teoretis penelitian ini terletak pada integrasi supervisi akademik, *Lesson Study*, kepemimpinan instruksional, PLC, dan CIPP dalam satu kerangka evaluasi yang menjelaskan hubungan sebab-akibat antarkomponen di konteks sekolah dasar; sedangkan kontribusi praktisnya berupa dasar pengambilan keputusan berbasis bukti bagi pemangku kepentingan agar kualitas pembelajaran di SDN 1 Marga bergerak seiring dengan penguatan refleksi guru dan kepemimpinan instruksional.

## METODE

### Desain Penelitian

Penelitian ini menggunakan pendekatan *evaluation research* dengan model CIPP (*Context, Input, Process, Product*) yang dikembangkan oleh Stufflebeam & Zhang (2017) dipadukan dengan pendekatan kuantitatif deskriptif. Model CIPP dipilih karena memberikan kerangka komprehensif untuk memotret siklus program secara utuh, mulai dari relevansi kebutuhan, kesiapan sumber daya, proses implementasi, hingga capaian hasil dan telah terbukti menghasilkan rekomendasi yang kontekstual dalam evaluasi program pendidikan. Data dianalisis secara deskriptif untuk menggambarkan kondisi objektif keempat komponen CIPP di SD Negeri 1 Marga tanpa menarik generalisasi kausal antara program dan hasil belajar, sehingga tetap selaras dengan tujuan evaluasi formatif dan sumatif program di satu konteks sekolah.

### Lokasi, Waktu, dan Subjek Penelitian

Penelitian dilaksanakan di SD Negeri 1 Marga, Kecamatan Marga, Kabupaten Tabanan, Bali, pada semester genap tahun pelajaran 2026/2027 (Juli - Agustus 2026). SD Negeri 1 Marga dipilih secara purposif karena sejak 2023 telah menerapkan program supervisi akademik berbasis *Lesson Study* secara bertahap dan tercatat dalam Rapor Pendidikan sebagai salah satu sekolah dengan peningkatan indikator proses (refleksi guru dan kepemimpinan instruksional) yang tidak sebanding dengan stagnasi indikator hasil (kualitas pembelajaran). Subjek

penelitian mencakup seluruh pemangku kepentingan utama program di sekolah: 1 pengawas pembina yang menangani SDN 1 Marga, 1 kepala sekolah, dan seluruh guru yang terlibat dalam program supervisi akademik berbasis *Lesson Study*. Mengingat ukuran populasi terjangkau, teknik pengambilan subjek menggunakan total sampling sehingga seluruh anggota populasi di SDN 1 Marga dijadikan responden untuk memaksimalkan keterwakilan perspektif dalam evaluasi.

### Variabel dan Definisi Operasional

Empat komponen CIPP dioperasionalkan sebagai berikut untuk konteks SDN 1 Marga:

#### 1. Context:

- Kesesuaian tujuan program supervisi akademik berbasis *Lesson Study* dengan kebutuhan pengembangan kompetensi guru di SDN 1 Marga.
- Kesesuaian program dengan kebijakan yang berlaku (misalnya Permendiknas No. 12 Tahun 2007 tentang Standar Pengawas Sekolah/Madrasah dan implementasi Kurikulum Merdeka di sekolah).
- Analisis kebutuhan supervisi dan *Lesson Study* di tingkat sekolah (misalnya kebutuhan pembinaan berdasarkan hasil Rapor Pendidikan dan hasil evaluasi kinerja guru).

#### 2. Input:

- Kompetensi supervisor (pengawas dan kepala sekolah) dalam merancang dan melaksanakan supervisi akademik berbasis *Lesson Study*.
- Kesiapan guru (motivasi, pengetahuan awal tentang *Lesson Study*, komitmen berpartisipasi).
- Ketersediaan sarana-prasarana pendukung (waktu kolaborasi, ruang pertemuan, perangkat pembelajaran).
- Ketersediaan dan kualitas pedoman program (SOP, jadwal, format rencana dan laporan *Lesson Study*).

#### 3. Process:

- Keterlaksanaan siklus *Lesson Study* (*Plan-Do-See*) dalam kerangka supervisi akademik: kualitas perencanaan bersama, pelaksanaan pembelajaran yang diobservasi, dan sesi refleksi.
- Peran supervisor (pengawas dan kepala sekolah) dalam memfasilitasi, memonitor, dan menindaklanjuti proses supervisi dan *Lesson Study*.
- Hambatan implementasi di tingkat sekolah (misalnya keterbatasan waktu, resistensi guru, kendala koordinasi).

#### 4. Product:

- Peningkatan kompetensi pedagogis dan profesional guru sebagaimana dipersepsikan oleh guru, kepala sekolah, dan pengawas.
- Keberlanjutan komunitas belajar profesional (PLC) di SDN 1 Marga setelah siklus *Lesson Study* berjalan beberapa tahun.
- Capaian kualitas pembelajaran di kelas yang diverifikasi secara triangulatif melalui kuesioner, dokumen Rapor Pendidikan (indikator D.1), dan observasi pembelajaran.

### Teknik dan Instrumen Pengumpulan Data

Data dikumpulkan melalui empat teknik utama yang saling melengkapi:

#### 1. Kuesioner

Kuesioner ber-skala *Likert* 1–5 digunakan sebagai instrumen utama untuk seluruh responden (pengawas, kepala sekolah, dan guru). Kisi-kisi kuesioner memuat 60 butir yang tersebar pada empat komponen CIPP (Context 10 butir, Input 14 butir, Process 20 butir, Product 16 butir). Butir disusun paralel untuk tiga kelompok responden dengan redaksi

disesuaikan peran masing-masing (misalnya butir tentang peran supervisi dirumuskan berbeda untuk pengawas, kepala sekolah, dan guru).

## 2. Wawancara Semi Terstruktur

Wawancara dilakukan terhadap informan yang dipilih secara purposif di SDN 1 Marga (1 pengawas, 1 kepala sekolah, dan beberapa guru yang aktif dalam *Lesson Study*) untuk menggali secara mendalam kendala, praktik baik, dan dinamika implementasi pada komponen *Process* dan *Product*. Panduan wawancara disusun selaras dengan indikator CIPP dan fokus pada penjelasan di balik temuan kuantitatif.

## 3. Observasi

Observasi dilakukan terhadap keterlaksanaan siklus *Plan-Do-See* di SDN 1 Marga menggunakan lembar observasi berskala 1–4 yang memotret kualitas perencanaan bersama, pelaksanaan pembelajaran yang diobservasi, dan sesi refleksi guru. Observasi juga digunakan untuk menguatkan penilaian terhadap aspek kepemimpinan instruksional dan PLC di sekolah.

## 4. Studi Dokumentasi

Dokumen yang dikaji meliputi: perencanaan program supervisi dan *Lesson Study*, presensi kegiatan, notulensi atau catatan refleksi, RPP atau modul ajar yang digunakan dalam siklus *Lesson Study*, serta Rapor Pendidikan tahun 2024–2025 untuk melihat tren indikator D.1, D.2, dan D.3 di SDN 1 Marga.

## Validitas dan Reliabilitas Instrumen

Validitas isi kuesioner diuji melalui penilaian pakar (*expert judgment*) menggunakan formula Gregory, yaitu:

$$\text{Validitas Isi} = \frac{D}{A + B + C + D}$$

Dimana A, B, C, dan D merepresentasikan kesesuaian dan ketidaksesuaian penilaian pakar terhadap butir instrumen. Validitas empiris diuji pada 30 responden di luar konteks SDN 1 Marga (misalnya guru dari sekolah lain yang memiliki karakteristik serupa) menggunakan korelasi *Product Moment Pearson*, dengan batas  $r$  tabel sebesar 0,361 untuk  $n = 30$  dan  $\alpha = 5\%$ . Butir dianggap valid apabila nilai  $r$  hitung  $\geq r$  tabel. Reliabilitas kuesioner dianalisis menggunakan koefisien Alpha Cronbach dengan bantuan perangkat lunak statistik; instrumen dinyatakan reliabel apabila koefisien  $\alpha \geq 0,70$ , yang menunjukkan konsistensi internal antarbutir pada masing-masing komponen CIPP.

## Teknik Analisis Data

Analisis data kuantitatif menggunakan statistik deskriptif yang dipadukan dengan analisis Kuadran Glickman untuk memetakan prototipe efektivitas program di SDN 1 Marga. Langkah analisis meliputi:

1. Menyusun tabulasi skor mentah per komponen CIPP untuk tiap responden.
2. Mengubah skor mentah menjadi Z-skor untuk menstandarkan distribusi.
3. Mengonversi Z-skor menjadi T-skor agar lebih mudah diinterpretasikan secara praktis.
4. Menentukan arah positif (+) atau negatif (–) untuk tiap komponen CIPP berdasarkan rata-rata T-skor terhadap titik potong yang ditetapkan, kemudian memetakan posisi program dalam kuadran efektivitas.

Penelitian ini mengadaptasi kerangka Glickman dengan menambahkan syarat bahwa komponen *Product* wajib bernilai positif agar program dikategorikan minimal Efektif. Penyesuaian metodologis ini dilakukan karena *Product* merepresentasikan hasil akhir program, bukan sekadar salah satu dari empat komponen yang setara. Kombinasi 16 prototipe CIPP dikelompokkan menjadi lima kategori:

- Sangat Efektif (4 komponen bernilai positif).
- Efektif (3 komponen bernilai positif termasuk Product).
- Cukup Efektif (1–2 komponen bernilai positif dengan Product tetap positif).
- Kurang Efektif (Product bernilai negatif berapa pun jumlah komponen lain yang positif).
- Tidak Efektif (tidak ada komponen yang bernilai positif).

Analisis dilakukan berlapis pada tiga level: (1) keseluruhan responden di SDN 1 Marga, (2) per komponen/indikator CIPP, dan (3) per kelompok responden (pengawas, kepala sekolah, guru), untuk mendeteksi variasi persepsi antarperan di dalam satu sekolah. Data kualitatif dari wawancara dianalisis secara tematik dan diintegrasikan melalui triangulasi metode (kuesioner, wawancara, observasi, dokumentasi) guna memperkuat interpretasi temuan kuantitatif serta menjelaskan faktor-faktor di balik kesenjangan proses–produk yang teridentifikasi.

### Kriteria Keberhasilan Program

Program supervisi akademik berbasis Lesson Study di SDN 1 Marga dinyatakan Efektif apabila memenuhi dua syarat utama:

1. Komponen Product bernilai positif secara konsisten.
2. Jumlah komponen CIPP yang bernilai positif  $\geq 3$  dari 4 komponen.

Jika komponen Product bernilai negatif, program dikategorikan Kurang Efektif meskipun tiga komponen lain bernilai positif. Penegasan ini dimaksudkan untuk menekankan bahwa proses yang kuat dan input yang memadai belum dapat dianggap berhasil apabila tidak menghasilkan perbaikan nyata pada kualitas pembelajaran di kelas sebagai tujuan utama program.

## HASIL DAN PEMBAHASAN

### HASIL

Penelitian ini melibatkan 10 responden di SD Negeri 1 Marga yang dipilih melalui teknik *total sampling*, terdiri atas 8 guru (guru kelas I–VI, guru PAH, dan guru PJOK), 1 kepala sekolah, dan 1 pengawas pembina Gugus 3. Bagian ini menyajikan temuan penelitian secara berurutan sesuai tiga tujuan yang dirumuskan pada bagian Pendahuluan: (1) efektivitas Program Supervisi Akademik Berbasis Lesson Study ditinjau dari empat komponen CIPP, (2) variasi persepsi antarperan dan pola individual yang berpotensi menjelaskan kesenjangan proses-produk, dan (3) temuan pendukung dari studi dokumentasi serta data Rapor Pendidikan sebagai dasar triangulasi.

### Efektivitas Program Ditinjau dari Empat Komponen CIPP

Secara umum, ketiga komponen (Konteks, Proses, dan Produk) berada pada kategori Baik, sedangkan komponen Masukan menjadi satu-satunya komponen yang jatuh ke kategori Cukup/Sedang. Konteks memperoleh mean tertinggi sekaligus variasi antarresponden terbesar, sementara Produk yang merepresentasikan hasil akhir program menunjukkan mean terendah kedua dengan rentang skor individu terlebar (3,38 hingga 4,06). Statistik deskriptif keempat komponen disajikan pada Tabel 1.

Tabel 1. Statistik Deskriptif Komponen CIPP di SDN 1 Marga (n = 10)

| Komponen | Mean | SD   | Min  | Max  | Kategori     |
|----------|------|------|------|------|--------------|
| Konteks  | 3,91 | 0,29 | 3,20 | 4,20 | Baik         |
| Masukan  | 3,58 | 0,28 | 3,14 | 3,93 | Cukup/Sedang |
| Proses   | 3,81 | 0,21 | 3,40 | 4,05 | Baik         |
| Produk   | 3,79 | 0,20 | 3,38 | 4,06 | Baik         |

Sumber: data Riset

## Variasi Persepsi Antarkelompok Responden

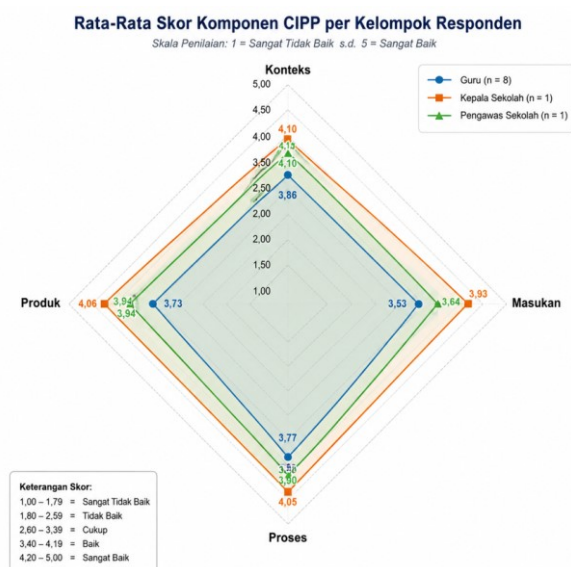
Ditinjau per kelompok peran, pola skor menunjukkan gradasi yang konsisten: Kepala Sekolah dan Pengawas Sekolah menilai program lebih tinggi dibanding Guru pada seluruh komponen, dengan selisih terbesar terjadi pada komponen Masukan antara Kepala Sekolah dan Guru. Guru sebagai kelompok terbanyak ( $n = 8$ ) menjadi satu-satunya kelompok yang skor Masukannya jatuh ke bawah kategori Baik. Perbandingan lengkap ketiga kelompok disajikan pada Tabel 2.

**Tabel 2. Rata-Rata Skor Komponen CIPP per Kelompok Responden**

| Kelompok         | n | Konteks | Masukan | Proses | Produk |
|------------------|---|---------|---------|--------|--------|
| Guru             | 8 | 3,86    | 3,53    | 3,77   | 3,73   |
| Kepala Sekolah   | 1 | 4,10    | 3,93    | 4,05   | 4,06   |
| Pengawas Sekolah | 1 | 4,10    | 3,64    | 3,90   | 3,94   |

Sumber: data Riset

Sebagai ilustrasi visual perbandingan pola antarkelompok pada keempat komponen sekaligus, data pada Tabel 2 dapat ditampilkan dalam diagram radar.



Sumber: Hasil Riset

**Gambar 1. Diagram Radar Perbandingan Skor CIPP Antarkelompok Responden**

## Pemetaan Individual Melalui Transformasi Z-Skor dan T-Skor

Setelah skor mentah distandardisasi menjadi Z-skor dan dikonversi menjadi T-skor terhadap titik potong  $T = 50$ , ditemukan rentang T-skor yang lebar antarresponden pada seluruh komponen, dengan satu responden (G6, Guru Kelas VI) secara konsisten berada pada T-skor terendah di keempat komponen (25,7–36,7), sedangkan Kepala Sekolah secara konsisten berada pada T-skor tertinggi (56,5–63,8). Anomali lain ditemukan pada G3 (Guru Kelas III), yang T-skor Proses-nya tinggi (59,0) namun T-skor Produk-nya justru di bawah titik potong (41,6) pola berlawanan arah yang tidak ditemukan pada responden lain. Hasil transformasi lengkap disajikan pada Tabel 3.

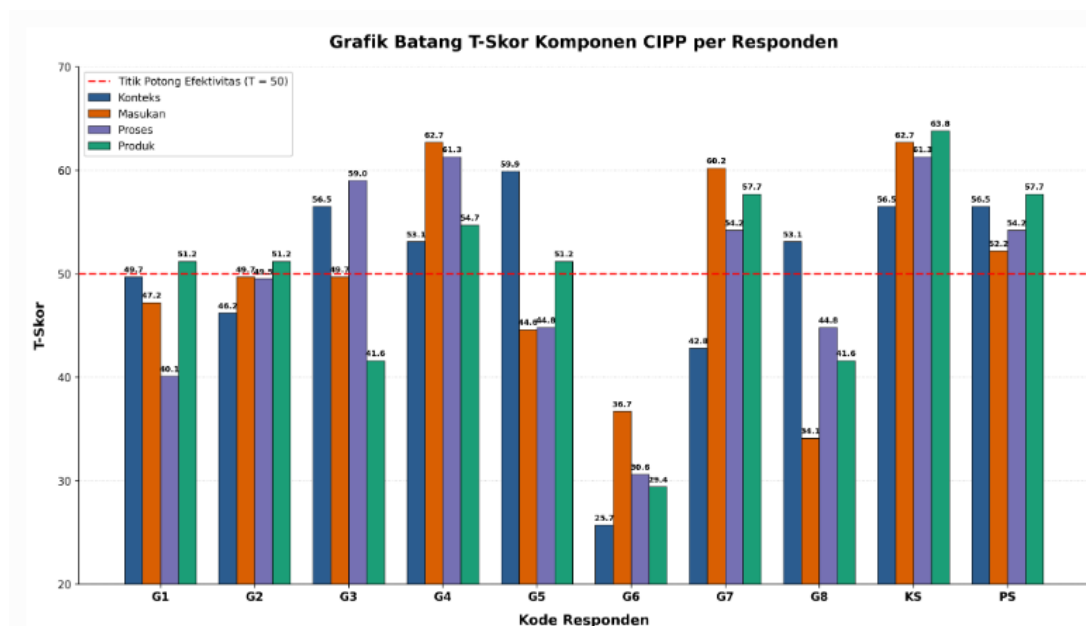
**Tabel 3. Hasil Transformasi T-Skor per Komponen CIPP per Responden**

| Kode | Peran          | Konteks (T) | Masukan (T) | Proses (T) | Produk (T) |
|------|----------------|-------------|-------------|------------|------------|
| G1   | Guru Kelas I   | 49,7        | 47,2        | 40,1       | 51,2       |
| G2   | Guru Kelas II  | 46,2        | 49,7        | 49,5       | 51,2       |
| G3   | Guru Kelas III | 56,5        | 49,7        | 59,0       | 41,6       |

|    |                  |      |      |      |      |
|----|------------------|------|------|------|------|
| G4 | Guru Kelas IV    | 53,1 | 62,7 | 61,3 | 54,7 |
| G5 | Guru Kelas V     | 59,9 | 44,6 | 44,8 | 51,2 |
| G6 | Guru Kelas VI    | 25,7 | 36,7 | 30,6 | 29,4 |
| G7 | Guru PAI         | 42,8 | 60,2 | 54,2 | 57,7 |
| G8 | Guru PJOK        | 53,1 | 34,1 | 44,8 | 41,6 |
| KS | Kepala Sekolah   | 56,5 | 62,7 | 61,3 | 63,8 |
| PS | Pengawas Sekolah | 56,5 | 52,2 | 54,2 | 57,7 |

Sumber: data Riset

Pola pada Tabel 3 divisualisasikan melalui grafik batang berkelompok (lihat Gambar 2) untuk memperjelas posisi setiap responden relatif terhadap titik potong T = 50.



Sumber: Hasil Riset

Gambar 2. Grafik Batang T-Skor Komponen CIPP per Responden

### Klasifikasi Kuadran Glickman dan Distribusi Kategori Efektivitas

Berdasarkan pola arah (+/-) empat komponen dengan syarat Produk wajib positif, distribusi kategori efektivitas program pada level individu adalah: Sangat Efektif (3 responden), Efektif (1 responden), Cukup Efektif (3 responden), dan Kurang Efektif (3 responden) tidak satu pun responden berada pada kategori Tidak Efektif. Temuan yang menonjol adalah tiga responden (G3, G6, G8) dikategorikan Kurang Efektif semata-mata karena T-skor Produk mereka berada di bawah titik potong, terlepas dari status komponen lain yang bervariasi (G3 dan G8 bahkan positif pada Konteks). Klasifikasi lengkap disajikan pada Tabel 4.

Tabel 4. Pola Kuadran Glickman dan Kategori Efektivitas per Responden

| Kode | Peran          | Pola (Konteks-Masukan-Proses-Produk) | Kategori Efektivitas |
|------|----------------|--------------------------------------|----------------------|
| G1   | Guru Kelas I   | - - - +                              | Cukup Efektif        |
| G2   | Guru Kelas II  | - - - +                              | Cukup Efektif        |
| G3   | Guru Kelas III | + - + -                              | Kurang Efektif       |
| G4   | Guru Kelas IV  | + + + +                              | Sangat Efektif       |
| G5   | Guru Kelas V   | + - - +                              | Cukup Efektif        |
| G6   | Guru Kelas VI  | - - - -                              | Kurang Efektif       |
| G7   | Guru PAH       | - + + +                              | Efektif              |
| G8   | Guru PJOK      | + - - -                              | Kurang Efektif       |
| KS   | Kepala Sekolah | + + + +                              | Sangat Efektif       |

|    |                  |      |                |
|----|------------------|------|----------------|
| PS | Pengawas Sekolah | ++++ | Sangat Efektif |
|----|------------------|------|----------------|

Sumber: data Riset

### Temuan Pendukung dari Studi Dokumentasi dan Data Rapor Pendidikan

Skor Rapor Pendidikan Kemendikdasmen 2025 mencatat SDN 1 Marga memperoleh D.1 (Kualitas Pembelajaran) sebesar 60,39 berkategori Sedang, sedangkan D.2 (Refleksi dan Perbaikan Pembelajaran oleh Guru) sebesar 77,08 dan D.3 (Kepemimpinan Instruksional) sebesar 78,79, keduanya berkategori Baik. Hasil studi dokumentasi internal menemukan bahwa presensi kegiatan Lesson Study tersedia untuk sebagian besar siklus dengan partisipasi guru rata-rata di atas 80%, tetapi notulensi refleksi (tahap *See*) belum konsisten mencantumkan rencana tindak lanjut secara eksplisit di setiap sesi, dan format modul ajar hasil Lesson Study antarguru belum seragam. Rekapitulasi temuan dokumentasi disajikan pada Tabel 5.

**Tabel 5. Rekapitulasi Temuan Studi Dokumentasi Program di SDN 1 Marga**

| Jenis Dokumen                          | Kelengkapan | Temuan Utama                                                                |
|----------------------------------------|-------------|-----------------------------------------------------------------------------|
| Perencanaan program supervisi akademik | Cukup       | Jadwal rinci per gugus belum sepenuhnya disosialisasikan ke seluruh guru    |
| Jadwal dan daftar hadir Lesson Study   | Cukup       | Partisipasi guru rata-rata di atas 80%; kehadiran pengawas relatif terbatas |
| Notulensi refleksi (tahap <i>See</i> ) | Kurang      | Rencana tindak lanjut belum konsisten dicatat secara eksplisit              |
| Modul ajar/RPP hasil Lesson Study      | Cukup       | Format pencatatan hasil diskusi <i>Plan</i> belum seragam antarguru         |
| Rapor Pendidikan 2024–2025             | Lengkap     | D.1 = 60,39 (Sedang); D.2 = 77,08 (Baik); D.3 = 78,79 (Baik)                |

Sumber: hasil pengisian pedoman studi dokumentasi (2026)

Penyajian data di atas merupakan seluruh temuan mentah yang diperlukan untuk menjawab tiga tujuan penelitian secara berurutan, tanpa interpretasi mekanisme sebab-akibat atau kaitan dengan teori, kedua hal tersebut akan dibahas secara khusus pada bagian Pembahasan dengan struktur R-C-I (*Results Interpretation, Comparison with Prior Studies, Implication*) sesuai temuan kunci: (a) kesenjangan Masukan sebagai titik lemah program, (b) kesenjangan persepsi struktural vs. pelaksana, dan (c) anomali individual pada pola kuadran Glickman.

## PEMBAHASAN

Bagian pembahasan ini dirancang untuk menginterpretasikan temuan empiris guna menjawab dua tujuan analitis utama penelitian secara eksplisit, yaitu mengidentifikasi faktor-faktor determinan di balik kesenjangan antara penguatan proses dan stagnasi kualitas pembelajaran di SDN 1 Marga, serta merumuskan implikasi teoritis maupun praktisnya bagi evaluasi program dan supervisi akademik. Untuk memberikan eksplanasi yang runtut dan komprehensif, pembahasan disajikan melalui sintesis terhadap tiga temuan kunci penelitian.

### Temuan Pertama: Komponen Masukan sebagai Titik Lemah Struktural Program

Data pada Tabel 1 dan Tabel 3 menunjukkan bahwa Masukan adalah satu-satunya komponen CIPP yang jatuh ke kategori Cukup/Sedang, dengan T-skor terendah justru dicatat oleh guru (G6 dan G8), bukan oleh unsur pimpinan. Pola ini mengindikasikan bahwa defisit program tidak terletak pada rancangan tujuan (Konteks) atau keterlaksanaan siklus (Proses), melainkan pada fondasi sumber daya yang menopang keduanya alokasi waktu kolaboratif, pedoman teknis tertulis, dan dukungan sarana diskusi.

Secara mekanistik, kondisi ini dapat dijelaskan oleh sifat Lesson Study yang menuntut waktu kolaboratif terjadwal secara eksplisit di luar jam mengajar reguler; ketika alokasi waktu tersebut tidak dilembagakan secara formal dalam kalender akademik sekolah, guru

menanggung beban ganda antara tugas mengajar dan tuntutan kolaborasi, yang pada akhirnya menurunkan kesiapan dan kualitas partisipasi mereka pada tahap *Plan*. Efek beban ganda ini lebih terasa oleh guru sebagai pelaksana harian dibanding oleh kepala sekolah atau pengawas yang berperan sebagai perancang dan pengawas program.

Temuan ini sejalan dengan kajian sistematis (Indrawan et al., 2024) yang menegaskan bahwa keberlanjutan dampak Lesson Study bersifat kontingen terhadap dukungan struktural dan budaya kolejial, bukan konsekuensi otomatis dari pelaksanaan siklus itu sendiri. Sejalan pula dengan Supriyati et al., (2024) yang menemukan bahwa penguatan supervisi akademik berbasis *Lesson Study* di madrasah hanya efektif ketika disertai penguatan sarana dan pendampingan berkelanjutan oleh kepala madrasah, serta Darsih et al., (2025) yang menegaskan pengembangan kompetensi pedagogis guru melalui *Lesson Study* bergantung pada ketersediaan waktu kolaboratif yang terstruktur secara formal. Namun demikian, temuan ini berlawanan dengan hasil Setyawati & Jaelani (2026) pada evaluasi CIPP program supervisi akademik di SD Negeri 3 Karang Bongkot, yang justru menemukan komponen Input sebagai komponen dengan skor tertinggi. Perbedaan ini dapat dijelaskan oleh perbedaan fase implementasi program: SDN 3 Karang Bongkot kemungkinan berada pada tahap awal peluncuran dengan dukungan sumber daya dan pelatihan yang masih baru, sedangkan SDN 1 Marga telah menjalankan program selama tiga tahun sehingga dukungan awal (pelatihan, insentif, pendanaan rintisan) cenderung terkikis tanpa pembaruan struktural sebuah dinamika penurunan input pascapeluncuran yang belum banyak dibahas dalam literatur CIPP jangka panjang.

Implikasi teoretisnya, temuan ini menegaskan bahwa komponen Input dalam model CIPP tidak bersifat statis sepanjang siklus program multitalun, melainkan memerlukan evaluasi berkala (*rolling input assessment*) agar dukungan struktural tidak menyusut seiring waktu. Implikasi praktisnya, pengawas dan kepala sekolah SDN 1 Marga perlu melembagakan kembali alokasi waktu kolaboratif dan pedoman teknis tertulis sebagai bagian dari strategi keberlanjutan program, bukan hanya sebagai perangkat pada fase peluncuran awal.

### **Temuan Kedua: Kesenjangan Persepsi antara Unsur Struktural dan Pelaksana**

Tabel 2 memperlihatkan bahwa kepala sekolah dan pengawas secara konsisten menilai program lebih tinggi pada seluruh komponen dibanding guru, dengan selisih terbesar pada komponen Masukan. Kesenjangan ini menunjukkan adanya perbedaan sudut pandang evaluatif antara pihak yang merancang program dan pihak yang menjalankannya di kelas.

Perbedaan ini dapat dijelaskan oleh akses informasi yang tidak simetris: kepala sekolah dan pengawas memiliki akses penuh terhadap dokumen perencanaan dan laporan capaian program, sehingga penilaian mereka cenderung berbasis *program-as-designed* (program sebagaimana direncanakan di atas kertas), sedangkan guru menilai berdasarkan *program-as-experienced* (program sebagaimana dijalani dalam praktik harian). Kesenjangan sudut pandang semacam ini dikenal sebagai *fidelity gap* dalam evaluasi program, di mana keberhasilan program di atas kertas tidak selalu terkonfirmasi pada level pengalaman langsung penerima manfaat utama.

Temuan ini konsisten dengan Ngatini et al., (2025) yang menunjukkan kepemimpinan instruksional dan supervisi akademik secara simultan berkontribusi besar terhadap kualitas pembelajaran ketika diukur dari sisi kebijakan dan tindakan supervisor, namun kontribusi tersebut idealnya dimediasi oleh persepsi guru sebagai penerima langsung dampak supervisi. Wahyuni et al., (2026) juga menegaskan bahwa kepemimpinan instruksional kepala sekolah dalam penguatan manajemen sekolah dasar kerap terukur kuat pada indikator manajerial-administratif, namun tidak otomatis tercermin secara setara pada persepsi guru terhadap dukungan pedagogis harian. Sejalan dengan itu, tinjauan literatur Egiska Lusita Putri & Sucitra (2023) mencatat bahwa efektivitas kepemimpinan instruksional lebih sering dinilai dari

perspektif kepala sekolah sendiri tanpa triangulasi memadai dari guru, sehingga rentan terhadap bias penilaian yang menguntungkan diri sendiri (*self-serving assessment*).

Implikasi metodologisnya, temuan ini menguatkan pentingnya desain analisis berlapis bukan sekadar skor agregat sekolah sebagaimana telah diadaptasi pada penelitian ini, karena skor agregat berpotensi menutupi kesenjangan persepsi yang justru menjadi sumber informasi diagnostik penting. Implikasi praktisnya, kepala sekolah dan pengawas SDN 1 Marga perlu membangun mekanisme umpan balik guru yang terstruktur dan rutin untuk memvalidasi persepsi mereka terhadap capaian program, bukan hanya mengandalkan laporan administratif satu arah.

### **Temuan Ketiga: Disosiasi antara Proses dan Produk pada Level Individu**

Temuan paling substantif dari pemetaan kuadran Glickman (Tabel 4) adalah munculnya pola disosiasi pada G3 (Guru Kelas III), yang mencatat T-skor Proses tertinggi kedua (59,0) namun T-skor Produk di bawah titik potong (41,6) sebuah pola yang tidak ditemukan pada responden lain dan sekaligus menjadi bukti empiris paling langsung atas kesenjangan proses-produk yang menjadi problem statement penelitian ini. Bersama G6 dan G8 yang konsisten rendah pada seluruh komponen, ketiga responden ini menyumbang 30 persen dari total guru yang dikategorikan Kurang Efektif murni karena lemahnya skor Produk.

Disosiasi pada G3 mengindikasikan bahwa partisipasi aktif seorang guru dalam siklus *Plan-Do-See* tidak secara otomatis terkonversi menjadi perubahan praktik mengajar yang dipersepsikan sendiri oleh guru tersebut. Kemungkinan mekanismenya adalah refleksi yang dilakukan bersifat performatif untuk memenuhi kewajiban administratif observasi tanpa disertai transfer pembelajaran yang mendalam ke praktik mengajar individual, sebuah kemungkinan yang diperkuat oleh temuan studi dokumentasi bahwa notulensi refleksi tahap *See* belum konsisten mencantumkan rencana tindak lanjut secara eksplisit di setiap sesi.

Temuan ini sejalan dengan Indrawan et al., (2024) yang menegaskan bahwa siklus Lesson Study tidak otomatis mengubah praktik mengajar tanpa komunitas belajar profesional yang cukup kuat untuk menindaklanjuti hasil refleksi. Namun demikian, temuan ini berlawanan dengan Astuti et al., (2026) yang melaporkan peningkatan kompetensi pedagogik guru sekolah dasar secara konsisten setelah penerapan Lesson Study tanpa disosiasi proses-produk yang berarti. Perbedaan ini kemungkinan disebabkan oleh perbedaan desain dan durasi program: penelitian Astuti et al., (2026) bersifat deskriptif-kualitatif pada satu siklus program yang relatif baru, sedangkan SDN 1 Marga telah menjalani program selama tiga tahun, sehingga potensi kejenuhan reflektif (*reflection fatigue*) atau rutinisasi seremonial pada guru tertentu lebih mungkin terjadi sebuah dinamika longitudinal yang jarang tertangkap pada studi *cross-sectional* jangka pendek. Pola serupa juga dilaporkan oleh Mutiara Rushita Adi et al., (2025) dalam evaluasi CIPP terhadap program supervisi akademik, yang menemukan bahwa komponen Produk paling rentan terhadap variasi individual dibanding tiga komponen lain, sehingga memperkuat argumen bahwa Produk memerlukan perhatian evaluatif khusus dan tidak dapat diasumsikan bergerak linear mengikuti kekuatan Proses. Pola asimetri Proses-Produk ini pula yang teridentifikasi pada konteks pendidikan tinggi internasional oleh Babadi et al., (2024) dalam evaluasi CIPP program kedokteran gigi di Iran, serta oleh Sankaran & Saad (2022) pada evaluasi program Sarjana Pendidikan, yang keduanya menemukan komponen Proses dinilai lebih tinggi dan lebih homogen dibanding komponen Produk mengindikasikan bahwa asimetri Proses-Produk bukan fenomena khas konteks sekolah dasar di Indonesia, melainkan pola umum dalam evaluasi program berbasis CIPP lintas jenjang dan lintas negara.

Implikasi teoretisnya, temuan ini memperkuat penegasan metodologis Stufflebeam & Zhang (2017) serta adaptasi kerangka Glickman pada penelitian ini bahwa komponen Produk harus dievaluasi secara independen dan tidak dapat diprediksi otomatis dari kekuatan komponen Proses. Kontribusi ini turut mengisi celah yang diidentifikasi Citariani et al., (2025)

dalam kajian sistematis mereka, yaitu minimnya studi CIPP yang menelusuri hubungan sebab-akibat antarkomponen secara mendalam pada satu konteks sekolah spesifik. Implikasi praktisnya, kepala sekolah dan pengawas SDN 1 Marga perlu merancang mekanisme pendampingan pascasiklus yang bersifat individual bukan hanya kelompok untuk memastikan hasil refleksi benar-benar diterapkan oleh setiap guru, khususnya bagi guru dengan pola disosiasi seperti G3.

Merujuk pada tiga tujuan penelitian yang dirumuskan pada bagian Pendahuluan, hasil dan pembahasan di atas memberikan tiga jawaban tegas. Pertama, terkait evaluasi efektivitas Program Supervisi Akademik Berbasis Lesson Study melalui model CIPP, program ini tergolong Baik hingga Sangat Efektif pada level agregat sekolah, tetapi tidak konsisten efektif pada level individu, 30 persen guru berada pada kategori Kurang Efektif karena lemahnya komponen Produk, sehingga klaim keberhasilan program tidak dapat digeneralisasi secara merata ke seluruh guru. Kedua, terkait faktor yang menjelaskan kesenjangan proses-produk, tiga mekanisme teridentifikasi: defisit struktural pada komponen Masukan yang terkikis seiring durasi program, kesenjangan persepsi antara unsur struktural (kepala sekolah, pengawas) dan pelaksana (guru) akibat *fidelity gap*, serta disosiasi individual antara partisipasi proses dan perubahan praktik mengajar akibat refleksi yang berpotensi performatif. Ketiga, terkait rekomendasi strategis, temuan ini mengarahkan pada kebutuhan evaluasi input berkala, mekanisme umpan balik guru yang terstruktur, dan pendampingan individual pasca siklus yang akan dielaborasi lebih lanjut pada bagian Kesimpulan.

Penelitian ini memiliki tiga keterbatasan utama yang perlu diakui secara jujur. Pertama, jumlah responden yang terbatas pada satu sekolah ( $n = 10$ ) membatasi generalisasi temuan ke sekolah lain di Gugus 3 Kecamatan Marga maupun ke konteks sekolah dasar secara umum, sehingga interpretasi hasil sebaiknya dipahami sebagai studi kasus mendalam, bukan representasi populasi yang lebih luas. Kedua, data kuesioner bersifat persepsi diri (*self-report*) yang berpotensi mengandung bias keinginan sosial (*social desirability bias*), khususnya pada responden dengan relasi hierarkis terhadap program yang dievaluasi (misalnya kepala sekolah yang turut merancang program). Ketiga, desain penelitian yang bersifat deskriptif *cross-sectional* tidak memungkinkan penarikan kesimpulan kausal antara kekuatan komponen Proses dan lemahnya komponen Produk pada level individu, sehingga penjelasan mekanistik mengenai disosiasi pada G3 dalam pembahasan ini bersifat hipotesis yang memerlukan konfirmasi lebih lanjut melalui triangulasi kualitatif yang lebih dalam.

Berdasarkan keterbatasan di atas, penelitian selanjutnya disarankan mengarah pada tiga hal konkret. Pertama, replikasi studi pada seluruh delapan sekolah di Gugus 3 Kecamatan Marga secara komparatif akan memungkinkan pengujian apakah pola defisit Masukan dan disosiasi Proses-Produk yang ditemukan di SDN 1 Marga bersifat spesifik-sekolah atau merupakan pola sistemik gugus. Kedua, desain longitudinal yang melacak indikator D.1 Rapor Pendidikan selama beberapa tahun berjalannya program akan memberikan bukti yang lebih kuat mengenai hipotesis kejenuhan reflektif (*reflection fatigue*) yang diajukan dalam pembahasan ini. Ketiga, penelitian kualitatif mendalam melalui wawancara dan observasi kelas langsung pada guru dengan pola disosiasi seperti G3 diperlukan untuk mengonfirmasi mekanisme di balik lemahnya konversi hasil refleksi Lesson Study menjadi perubahan praktik mengajar yang terukur.

## KESIMPULAN

Penelitian ini mengevaluasi Program Supervisi Akademik Berbasis Lesson Study di SD Negeri 1 Marga Kabupaten Tabanan melalui model CIPP dan menemukan bahwa program tersebut tergolong Baik hingga Sangat Efektif pada level agregat sekolah, tetapi efektivitas ini tidak terdistribusi secara merata pada level individu guru. Komponen Konteks, Proses, dan Produk berada pada kategori Baik, sedangkan komponen Masukan menjadi satu-satunya

komponen yang jatuh ke kategori Cukup/Sedang, dan tiga dari delapan guru dikategorikan Kurang Efektif semata-mata karena skor Produk mereka berada di bawah titik potong efektivitas termasuk satu guru yang justru mencatat keterlibatan proses tertinggi kedua di sekolah. Temuan ini menjawab tujuan pertama penelitian secara tegas: peningkatan pada dimensi proses dan kepemimpinan instruksional di SDN 1 Marga belum secara otomatis dan merata terkonversi menjadi perbaikan kualitas pembelajaran di kelas, sehingga kesenjangan proses-produk yang menjadi problem statement penelitian ini terkonfirmasi secara empiris pada level sekolah tunggal, bukan sekadar pola agregat gugus.

Terkait tujuan kedua, penelitian ini mengidentifikasi tiga faktor yang menjelaskan kesenjangan tersebut, yaitu defisit struktural pada dukungan sumber daya dan waktu kolaboratif yang cenderung menyusut sejalan dengan durasi program multitalahun, kesenjangan sudut pandang antara unsur struktural yang menilai program berdasarkan desain dan unsur pelaksana yang menilai berdasarkan pengalaman langsung di kelas, serta disosiasi pada level individu antara partisipasi dalam siklus Plan-Do-See dan perubahan praktik mengajar yang dipersepsikan sendiri oleh guru. Ketiga faktor ini secara bersama-sama menunjukkan bahwa keberhasilan Lesson Study dan supervisi akademik tidak dapat diasumsikan berjalan linear dari kekuatan proses menuju hasil, melainkan memerlukan penguatan yang disengaja pada dimensi input dan mekanisme pendampingan individual pascarefleks. Berkenaan dengan tujuan ketiga, temuan ini memberi dasar bagi pengawas dan kepala sekolah SDN 1 Marga untuk memprioritaskan evaluasi input secara berkala, membangun kanal umpan balik guru yang terstruktur untuk memvalidasi persepsi pimpinan terhadap capaian program, dan merancang pendampingan reflektif yang bersifat individual bagi guru dengan pola disosiasi proses-produk, sebagai langkah lebih tepat sasaran dibanding penghentian atau perluasan program secara seragam.

Secara teoretis, penelitian ini memberikan kontribusi bagi bidang evaluasi program pendidikan dengan menunjukkan bahwa dalam kerangka CIPP, komponen Produk perlu diperlakukan sebagai unit evaluasi yang independen dan tidak dapat diprediksi otomatis dari kekuatan komponen Proses, sekaligus memperkuat kebutuhan analisis berlapis individu, kelompok peran, dan sekolah dalam mengevaluasi program pengembangan profesional guru pada satu satuan pendidikan. Kontribusi ini memperkaya penerapan model evaluasi CIPP pada konteks sekolah dasar dengan data empiris yang secara spesifik menunjukkan mekanisme di balik kesenjangan proses-produk, sebuah aspek yang selama ini belum banyak ditelusuri secara mendalam pada level satu sekolah. Mengingat penelitian ini terbatas pada satu sekolah dengan sepuluh responden, temuan ini perlu dimaknai sebagai bukti kontekstual yang mendalam pada kasus SDN 1 Marga, bukan sebagai generalisasi bagi seluruh sekolah dasar di Kecamatan Marga atau di Kabupaten Tabanan secara keseluruhan.

## REFERENSI

- Astuti, D., Wasliman, E. D., & Handayani, S. (2026). *Strategi Peningkatan Kompetensi Pedagogik Guru Berbasis Lesson Study di Sekolah Dasar*. 9, 1954–1963.
- Azizah, K., & Nadia, H. (2024). *Evaluasi Kinerja Guru Melalui Lesson Study di Sekolah dasar Wilayah Telukjambe Timur*. 2(1), 108–123.
- Babadi, F., Esfandiari, M., & Cheraghi, M. (2024). Evaluating the dentistry program in Iran using the context, input, process, and product (CIPP) model: a comprehensive analysis. *Frontiers in Medicine*, 11(July), 1–6. <https://doi.org/10.3389/fmed.2024.1394395>
- Citariani, N. M., Santi, I. K. L., & Riawan, I. M. O. (2025). Efektivitas Model Evaluasi CIPP dalam Program Pendidikan: Kajian Literatur Sistematis. *Jurnal Evaluasi Pendidikan*, 16, 108–115. <https://journal.unj.ac.id/unj/index.php/jep/article/view/58457>
- Darsih, E., Hanggara, A., & Putra, R. A. (2025). *Leveraging Lesson Study to Develop Technological Pedagogical Content Knowledge in Pre-Service Teacher Education*.

- 8(3), 611–622.
- Egiska Lusita Putri, & Sucitra, D. A. (2023). Kepemimpinan Instruksional Kepala Sekolah Literature Review Egiska. *Jurnal Manajemen Mutu Pendidikan*, 11(2), 47–55. <https://doi.org/10.23960/jmmp>
- Ichwan, R., & Sukandar, A. (2026). *EVALUASI PROGRAM PENGEMBANGAN PROFESIONAL GURU MELALUI SUPERVISI AKADEMIK DI ERA REVOLUSI INDUSTRI 5.0*. 11(2), 1983–1992.
- Indrawan, I. M., Sumarni, N. W., & Suarni, K. N. (2024). Lesson study implementation and its challenges: A systematic review. *Jurnal Pendidikan Indonesia*, 13(1), 45–58. <https://doi.org/https://doi.org/10.23887/jpi-undiksha.v13i1.71453>
- Lewis, C. (2016). *How does Lesson Study Transform Teaching and Teachers' Learning?*
- Mutiara Rushita Adi, L. P., Hendra Divayana, D. G., & Gede Agung, A. A. (2025). Penerapan Model Evaluasi CIPP dalam meningkatkan Efektivitas Program Supervisi Akademik. *Jurnal Manajemen Pendidikan Dan Ilmu Sosial*, 6(1), 643–652. <https://doi.org/10.38035/jmpis.v6i1.3340>
- Ngatini, Kusumaningsih, W., Nurkolis, & Baldemor, M. R. (2025). Influence of Instructional Leadership and Academic Supervision on The Quality of Learning. *IJORER : International Journal of Recent Educational Research*, 6(2), 602–614. <https://doi.org/10.46245/ijorer.v6i2.804>
- Permendikdasmen No. 1 (2025).
- Salasiah, S., Haryono, A., & Supriyati, W. (2023). Pengembangan Perangkat Pembelajaran Ipa Berbasis Lesson Study Untuk Meningkatkan Kolaborasi Guru Dan Kemampuan Berpikir Kritis Siswa. *Binomial*, 6(1), 58–71. <https://doi.org/10.46918/bn.v6i1.1728>
- Sankaran, S., & Saad, N. (2022). Evaluating the Bachelor of Education Program Based on the Context, Input, Process, and Product Model. *Frontiers in Education*, 7(June), 1–8. <https://doi.org/10.3389/educ.2022.924374>
- Setyawati, E., & Jaelani, A. K. (2026). *Evaluasi Program Pelaksanaan Supervisi Akademik Menggunakan Model CIPP di SD Negeri 3 Karang Bongkot , Kecamatan Labuapi*. 10(1). <https://doi.org/10.29303/jpap.v10i1.1349>
- Stufflebeam, D. L., & Zhang, G. (2017). The CIPP evaluation model: How to evaluate for improvement and accountability. In *Guilford Press*. <https://doi.org/https://doi.org/10.1007/978-3-319-64529-3>
- Sundari, T. (2023). Peningkatan Mutu Pembelajaran Melalui Supervisi AKademik Berbasis Lesson Study. *Equity in Education Journal (EEJ)*, 6(1), 46–53. <https://e-journal.upr.ac.id/index.php/eej/article/view/2447/2221>
- Supriyati, N., Baharuddin, B., & Sutiah, S. (2024). Enhancing Academic Supervision of Madrasah Principals through Lesson Study Initiatives. *AL-ISHLAH: Jurnal Pendidikan*, 16(4), 5325–5334. <https://doi.org/10.35445/alishlah.v16i4.4965>
- Ulum, A. P. B. (2025). Jurnal Pendiidkan Indonesia : IMPLEMENTASI PROGRAM PENGEMBANGAN PROFESIONAL. *Jurnal Pendiidkan Indonesia*, 5(5). <https://doi.org/10.59818/jpi.v5i5.2157>
- Wahyuni, D. D., Novita, R., Detri, Morelent, Y., & Satria, E. (2026). *Kepemimpinan Instruksional Kepala Sekolah dalam Penguatan Manajemen Sekolah Dasar*. 12, 306–312.