

# EVALUATING SCHOOL QUALITY THROUGH AN ADAPTED CONTEXT-PROCESS-PRODUCT FRAMEWORK

Zainul Arifin<sup>1\*</sup>, Izakk Hendrik Wenno<sup>2</sup>, Emma Rumahlewang<sup>3</sup>

<sup>123</sup>Universitas Pattimura, Indonesia

\*Corresponding Author: zinul.arif@gmail.com

## ABSTRACT

School quality improvement requires attention to both internal organizational processes and external governance. This study evaluated a school quality program through an adapted application of the Context, Input, Process, and Product framework, operationalizing Context as school committee role, Process as principal adaptive leadership, and Product as school quality culture because Input data were unavailable. A descriptive-correlational design involved 18 voluntary teachers and education personnel from one school. Three 0-4 questionnaires measured school committee roles (40 items), adaptive leadership (36 items), and school quality culture (36 items). Aggregate attainment was 81.6% for Context, 89.7% for Process, and 88.9% for Product. Two-tailed Pearson analysis identified a positive Process-Product association ( $r = 0.551$ ,  $p = 0.018$ ; approximate 95% CI [0.113, 0.810]), whereas Context was not significantly associated with Process or Product. The five Context items with the largest standard deviations all concerned committee oversight, indicating substantial response heterogeneity. Eight respondents could not be matched to the available official teacher roster, and instrument provenance, psychometric evidence, and detailed ethics documentation were unavailable in the study record. The findings are therefore exploratory and site-specific. The study contributes a cautious, CIPP-informed diagnostic perspective that distinguishes internal leadership-quality culture alignment from perceived external oversight and identifies priorities for stronger governance documentation and future mixed-method evaluation.

**Keywords:** adaptive leadership, program evaluation, school committee, school governance, school quality culture

## INTRODUCTION

School quality improvement is increasingly understood as an organizational and governance problem rather than a matter of isolated instructional interventions. Leadership matters because school leaders influence the conditions in which teachers collaborate, learn, and sustain improvement, although the strength and pathway of that influence vary across systems and contexts (Leithwood et al., 2020; Tan et al., 2024). In Indonesia, the literature similarly portrays school leadership as context-sensitive and closely tied to organizational culture, professional relationships, and the institutional environment in which schools operate (Lumban Gaol, 2023). This broader view is particularly relevant to basic education, where principals, teachers, education personnel, families, school committees, and public authorities share different parts of responsibility for quality, accountability, and improvement.

Program evaluation provides a structured way to connect these multiple dimensions. The CIPP framework - Context, Input, Process, and Product - was developed as a decision-oriented model intended to support systematic evaluation across program needs, resources, implementation, and outcomes (Stufflebeam, 2003). Historical analyses emphasize that CIPP evolved as a comprehensive evaluation architecture rather than as a simple causal model linking four variables in a linear sequence (Dizon, 2023). Contemporary applications demonstrate its flexibility across educational settings, including curriculum and program evaluation, while systematic reviews also show that the value of CIPP depends on the quality of evidence used to operationalize each component (Rocha et al., 2022). In Indonesia, CIPP-informed evaluations have been applied to teacher education and school programs (Cahyadi et al., 2022; Kasmaini et al., 2023; Kumalasari & Idawati, 2023; Sari et al., 2023), reinforcing its practical utility while also illustrating that component definitions must be explicitly justified for each evaluation context.

The Context dimension is especially pertinent when school quality is viewed through governance. Accountability research distinguishes between formal structures and the actual processes through which stakeholders obtain information, participate in decisions, and hold institutions responsible for performance. Public accountability frameworks caution that accountability mechanisms can fail when roles, information flows, and decision rights are unclear (Hutt & Polikoff, 2020). A systematic review by Ordofa and Asgedom (2022) likewise indicates that the relationship between school accountability and learning outcomes is contingent on how accountability is designed and enacted. Cross-national evidence further suggests that school governance arrangements interact with institutional conditions rather than producing uniform effects across countries (Luschei & Jeong, 2021). Studies of decentralization similarly show that transferring authority without corresponding clarity, capacity, and answerability can generate uneven accountability practices (Matete, 2022).

Stakeholder participation is therefore not equivalent to effective governance. School autonomy and accountability can create opportunities for improvement, but their effects depend on organizational capacity and on whether responsibilities are translated into transparent routines (Alajmi, 2022). Financial governance studies have also highlighted the importance of clear accountability practices at school level (Dwangu & Mahlangu, 2021), while research on intelligent accountability underscores the value of professional judgment, evidence use, and dialogue rather than compliance alone (Tolo et al., 2020). Parent and citizen participation can be formally embedded in school governance yet still vary in influence, voice, and recognition (Gerrard & Savage, 2022). Similarly, learner-participation research demonstrates that representation can become tokenistic when participation does not translate into meaningful involvement in governance (Yu & Shay, 2022). These findings justify examining not merely whether a school committee exists, but whether its functions are consistently visible and understood by staff.

This issue has direct relevance in Indonesia. Regulation of the Minister of Education and Culture of the Republic of Indonesia Number 75 of 2016 assigns school committees advisory, supportive, supervisory, and mediating functions. Empirical evidence from Indonesian schools shows that the quality of education services can be strengthened when committee roles are substantively implemented and communicated (Nuraya et al., 2021). Research on good school governance in Indonesian vocational schools similarly emphasizes transparency, accountability, participation, and decision quality (Supriadi et al., 2021). The gap between formal governance prescriptions and staff perceptions of actual oversight is therefore an important evaluative issue. A high overall score for committee role may coexist with substantial disagreement about specific oversight activities, and that disagreement can itself signal the need for closer diagnostic inquiry without automatically proving governance failure.

The Process dimension in the present study is represented by principal adaptive leadership. Adaptive leadership focuses on diagnosing challenges that cannot be solved through routine technical responses, mobilizing stakeholders, learning from changing conditions, and distributing responsibility for adaptation (Heifetz & Linsky, 2002). Contemporary school leadership research supports the relevance of collaborative and distributed practices to organizational improvement. Amels et al. (2021) found that the use of distributed leadership was related to teachers' capacity to change, while Liu et al. (2021) showed that supportive school culture and teacher collaboration can mediate relationships between leadership practices and teacher outcomes. A recent systematic review also shows that distributed school leadership is a multidimensional field in which leadership is understood as an organizational practice extending beyond the formal principal role (Mifsud, 2024). In Indonesian primary-school research, adaptive principal leadership has likewise been linked to teacher performance through collaborative school culture (Sumiati et al., 2024). These studies provide a plausible basis for

expecting leadership perceptions to be associated with perceptions of school quality culture, while not implying causality in a cross-sectional design.

Product is operationalized here as school quality culture, understood as the degree to which quality-oriented values and practices appear embedded in organizational routines. The concept builds on the broader total-quality tradition in education, which emphasizes continuous improvement, participation, and shared responsibility for quality (Sallis, 2002). More recent quality-culture research stresses that sustainable improvement depends on interaction among leadership support, ownership, communication, and collective learning rather than on isolated quality procedures (Bendermacher et al., 2021). School-based research similarly indicates that leadership can motivate improvement when members experience ownership of quality initiatives (Kirk & Barblett, 2022). Organizational culture and leadership are mutually embedded, with cultural expectations shaping leadership practice and leadership in turn influencing organizational norms (Bush, 2021). Empirical work on school culture also suggests that culture types are associated with leadership capacity (Demirhan, 2023). These perspectives support treating quality culture as an organizational condition relevant to the Product side of this adapted evaluation, while acknowledging that it is a perception-based proxy rather than a direct measure of student outcomes or institutional performance.

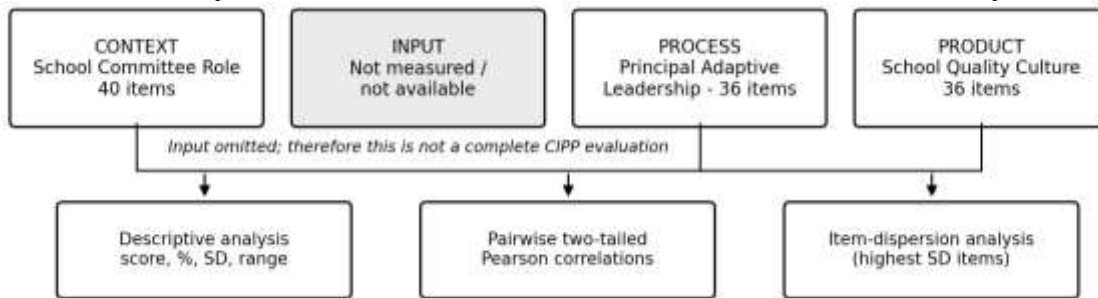
Despite these bodies of research, an empirical gap remains at the intersection of external school governance, internal adaptive leadership, and quality culture within a single diagnostic evaluation. Many CIPP studies retain the four canonical components and emphasize descriptive judgments about program implementation, whereas governance and leadership studies are often analyzed in separate literatures. Less attention has been given to whether a governance-related Context proxy covaries with internal leadership and quality-culture measures and whether staff perceptions are equally consistent across the specific functions of school oversight. The present study addresses this gap through a deliberately adapted three-component CIPP-informed structure and an item-dispersion analysis. Its novelty lies not in claiming a new version of CIPP, but in using the framework cautiously to combine three forms of diagnostic evidence: descriptive attainment, pairwise association, and within-item heterogeneity, while explicitly documenting how an initially very small respondent subset produced unstable interpretations.

Accordingly, this study aims to evaluate the school quality program by operationalizing school committee role as Context, principal adaptive leadership as Process, and school quality culture as Product; to examine the statistical relationships among these three study-specific components; and to identify school-committee items with the greatest response dispersion. Because the Input component was not measured, the study does not claim to evaluate the complete CIPP framework. The intended contribution is a site-specific governance and quality-improvement diagnosis that can inform local reflection and provide testable propositions for more rigorous multi-school and mixed-method research.

## **METHOD**

This study used a descriptive-correlational design and an adapted, rather than complete, CIPP-informed evaluation structure. The adaptation was conceptual and pragmatic. School committee role was operationalized as Context because it represents the external stakeholder and governance environment surrounding school management. Principal adaptive leadership was operationalized as Process because it reflects day-to-day leadership practices for managing change, participation, and organizational learning. School quality culture was operationalized as Product because it represents an organizational condition associated with the institutionalization of quality-oriented values and practices. Input was omitted because no corresponding Input measure was available in the study record. Consistent with the decision-

oriented nature of CIPP (Dizon, 2023; Stufflebeam, 2003), these mappings are treated as study-specific evaluation proxies rather than as a causal test of the canonical four-component model.



**Figure 1. Adapted CIPP-informed analytical structure used in the study**

The available study record identified 18 voluntary respondents comprising teachers and education personnel. All respondents worked in the same school. A probabilistic sampling procedure, a complete invitation log, and detailed eligibility criteria beyond membership in these staff categories were not documented. Accordingly, the sample is best understood as a small voluntary single-site respondent group rather than a probability sample. Cross-checking the respondent information against the available official teacher roster matched 10 respondents; eight respondents could not be matched to that roster. The available documentation did not permit verification of whether the unmatched respondents were education personnel, temporary staff, other eligible workers, or individuals omitted from an outdated roster. All respondents were therefore analyzed together, and no subgroup analysis by role, employment status, or tenure was conducted. Names are not reproduced in the manuscript in order to avoid unnecessary disclosure of identifiable information.

Data were collected through three questionnaires using a 0-4 response scale, where 0 represented Never/Disagree and 4 represented Always/Strongly Agree. The School Committee Roles questionnaire contained 40 items, the Principal Adaptive Leadership questionnaire contained 36 items, and the School Quality Culture questionnaire contained 36 items. The available documentation does not state whether these instruments were adopted, adapted, or newly developed, and does not provide supporting evidence for content validity, construct validity, pilot testing, or internal-consistency reliability. No psychometric properties are therefore asserted. This limitation is consequential because the substantive findings rely on respondents' scores on these measures; results are interpreted as exploratory perceptions from the available instruments rather than as estimates obtained from fully validated scales.

Participation was reported as voluntary. The available study documentation did not include a formal research-ethics approval number or sufficiently detailed records concerning written informed consent, confidentiality procedures, data retention, institutional permission, or data-access controls. The study therefore does not claim formal ethical approval that cannot be verified. In the revised report, findings are presented in aggregate form, respondent names are omitted, and the absence of fuller ethics and data-governance documentation is explicitly treated as a reporting limitation. Future replications should document institutional permission, participant information and consent, anonymization or pseudonymization procedures, storage arrangements, access controls, and retention or deletion rules before data collection begins.

Descriptive analysis used observed aggregate scores, maximum possible aggregate scores, percentages of the maximum possible score, respondent-level score ranges, and standard deviations. The percentage for each component was calculated as the observed aggregate score divided by its maximum possible aggregate score and multiplied by 100. In this article, the descriptor high is used only as a study-specific shorthand for an aggregate attainment above

80% of the maximum possible score; it is not presented as an externally validated cut-off. The standard deviations reported in Table 1 refer to respondents' raw total scores on each questionnaire rather than to percentage scores. Because the three instruments contain different numbers of items, dispersion should be interpreted relative to each scale rather than compared as if the raw standard deviations were standardized metrics.

Pairwise relationships among Context, Process, and Product total scores were evaluated using two-tailed Pearson correlations. Approximate 95% confidence intervals for the reported correlation coefficients were computed using the Fisher z transformation with  $n = 18$ . Respondent-level raw data were not available in the revision record, so assumptions relevant to Pearson correlation - including linearity, influential outliers, and distributional suitability - could not be re-examined independently. Inferential findings are therefore treated as exploratory and interpreted primarily through effect direction, magnitude, p values, and confidence intervals rather than as definitive population estimates. At the item level, the five statements with the largest standard deviations were identified as the items with the greatest response dispersion. Because frequencies by response category were unavailable, high dispersion is not equated with polarization or with a specific pattern of Never, Rarely, or Always responses.

A methodological sensitivity issue arose because an initial analysis had been conducted on a much smaller subset of six respondents before the respondent set was expanded to 18. In the  $n = 6$  subset, several school-committee oversight items appeared uniformly low; after expansion to  $n = 18$ , Context attainment increased and between-respondent variability became much larger. The smaller analysis is retained only as a cautionary illustration of instability in pattern-based interpretation from extremely small samples and is not used to support the final substantive conclusions. The final interpretation relies on the  $n = 18$  results reported below.

## RESULTS AND DISCUSSION

### Descriptive Profile of the Adapted Components

The descriptive results show high aggregate attainment across the three operationalized components, while the differences in score ranges and dispersion require careful interpretation. Table 1 reports the observed aggregate scores relative to the theoretical maximum for each instrument. Because Context contains 40 items and Process and Product contain 36 items each, raw totals are not directly comparable without reference to the corresponding maximum scores. The percentage column provides a normalized descriptive view of attainment, whereas the standard deviation retains the unit of respondent-level raw total score.

**Table 1. Summary of School Quality Program Scores Using the Adapted CIPP Components ( $n = 18$ )**

| Adapted Component                       | Total Score | Max Score | Average (%) | SD    | Range   |
|-----------------------------------------|-------------|-----------|-------------|-------|---------|
| Context - School Committee Role         | 2,351       | 2,880     | 81.6%       | 13.84 | 107-146 |
| Process - Principal Adaptive Leadership | 2,325       | 2,592     | 89.7%       | 5.62  | 123-144 |
| Product - School Quality Culture        | 2,305       | 2,592     | 88.9%       | 6.33  | 119-144 |

Context reached 81.6% of its maximum possible aggregate score, Process 89.7%, and Product 88.9%. These figures indicate that respondents tended to rate all three domains positively under the study-specific scoring convention. Process and Product were descriptively close, whereas Context was approximately eight percentage points lower than Process and seven points lower than Product. Context also showed the largest raw-score standard deviation

(13.84), compared with 5.62 for Process and 6.33 for Product. This should not be interpreted as a standardized comparison of construct variability because the instruments differ in length and lack documented psychometric calibration. At most, the result suggests that responses concerning the school committee role were less homogeneous within the available sample than responses concerning adaptive leadership and school quality culture.

### Associations Among the Adapted Components

The pairwise correlation results are presented in Table 2. Only the association between Process and Product met the conventional  $\alpha = .05$  threshold. Context showed negative coefficients with both Process and Product, but neither estimate was statistically significant and both confidence intervals included zero. Given the small sample, these non-significant directions should not be interpreted as evidence that stronger committee involvement reduces leadership quality or school quality culture.

**Table 2. Pairwise Two-Tailed Pearson Correlations Among the Adapted CIPP Components (n = 18)**

| Component Pair    | r      | p     | Approx. 95% CI  | Significant ( $\alpha = .05$ ) |
|-------------------|--------|-------|-----------------|--------------------------------|
| Context - Process | -0.325 | 0.188 | [-0.688, 0.167] | No                             |
| Context - Product | -0.357 | 0.146 | [-0.706, 0.132] | No                             |
| Process - Product | 0.551  | 0.018 | [0.113, 0.810]  | Yes                            |

The Process-Product coefficient ( $r = 0.551$ ,  $p = 0.018$ ) represents a moderate positive association in this respondent group. This pattern is theoretically plausible when viewed alongside international evidence showing that leadership practices can shape the conditions in which teachers collaborate and organizational improvement is sustained. Leithwood et al. (2020) describe leadership effects as operating through multiple organizational pathways rather than through direct effects alone, and Tan et al. (2024) similarly show through meta-analytic evidence that school leadership practices are associated with student and organizational outcomes through complex mechanisms. Liu et al. (2021) demonstrate that supportive culture and teacher collaboration can mediate relationships involving school leadership, while Amels et al. (2021) connect distributed leadership with teachers' capacity to change. In the Indonesian context, Sumiati et al. (2024) report a relationship between adaptive principal leadership and teacher performance through collaborative school culture. Taken together, these studies support the plausibility of a leadership-culture association without converting the present correlation into a causal claim.

The result also aligns conceptually with quality-culture scholarship. Sustainable quality development tends to depend on collective ownership, communication, leadership support, and continuous learning (Bendermacher et al., 2021). Kirk and Barblett (2022) show that leadership can support improvement initiatives through psychological ownership, while Demirhan (2023) links school culture types to school leadership capacity. Bush (2021) emphasizes that leadership and culture are reciprocally embedded in organizational and societal contexts. These perspectives help explain why respondents who perceive leadership as more adaptive may also perceive stronger quality-oriented routines. However, the same-source design is a major interpretive constraint: the same individuals evaluated both the principal and the school culture, so common-method variance, generalized satisfaction, or response style could inflate the observed association. Longitudinal, multi-informant, and observational data would be required to determine whether adaptive leadership precedes changes in quality culture.

The absence of statistically significant Context associations should not be described as a contradiction of CIPP. CIPP is an evaluation framework for organizing evidence about needs, resources, implementation, and outcomes, not a theory requiring positive pairwise correlations among operationalized component scores (Dizon, 2023; Stufflebeam, 2003). Reviews and contemporary CIPP applications show that evaluative conclusions depend on how each component is defined and on the sufficiency of the evidence collected for it (Cahyadi et al., 2022; Rocha et al., 2022). In the present study, Context is a perception-based proxy for the school committee role, Input is absent, Process is a leadership construct, and Product is an organizational culture construct. A non-significant correlation between these proxies can therefore reflect measurement characteristics, role distance, information asymmetry, sample uncertainty, or genuine organizational separation; the current data cannot discriminate among these explanations.

Governance literature reinforces the need for this cautious interpretation. Cross-national evidence suggests that school governance does not operate identically across systems and that associations with educational outcomes depend on institutional conditions (Luschei & Jeong, 2021). Accountability arrangements also vary according to how responsibilities, information, and authority are distributed (Hutt & Polikoff, 2020; Ordofa & Asgedom, 2022). Decentralization can produce uneven accountability when transferred responsibilities are not matched by capacity or clear answerability (Matete, 2022). In school-level studies, autonomy and accountability outcomes are similarly conditioned by implementation context (Alajmi, 2022), and financial management research underscores that formal responsibility must be accompanied by transparent practices (Dwangu & Mahlangu, 2021). Thus, the present non-significant Context coefficients are best treated as a local signal that perceived committee role did not covary reliably with internal leadership and culture perceptions in this sample, not as evidence that committee governance is irrelevant.

### **Committee-Oversight Response Dispersion**

Item-level analysis identified the five school-committee statements with the largest standard deviations. All five concern oversight: monitoring educational services, recommendations following supervision, oversight of the school budget, access to financial-management information, and periodic monitoring of school programs. Table 3 summarizes their means and standard deviations. Because category frequencies were not available, the analysis is intentionally limited to heterogeneity in ratings rather than a claim that the responses were polarized or that particular proportions selected low or high categories.

The five standard deviations range from 1.01 to 1.05 on a 0-4 scale. The defensible conclusion is that these items elicited comparatively dispersed judgments among respondents. The finding is substantively relevant because oversight is one of the formal functions expected of school governance structures. Indonesian evidence indicates that school committee effectiveness depends on substantive implementation of committee roles rather than formal existence alone (Nuraya et al., 2021), and good-governance research emphasizes transparency and participation in decision processes (Supriadi et al., 2021). Internationally, accountability studies likewise stress that access to information and clarity of responsibility are central to meaningful oversight (Hutt & Polikoff, 2020; Tolo et al., 2020). Yet dispersion by itself cannot establish that information was withheld, that the committee was inactive, or that the oversight mechanism failed. It may instead reflect differences in staff role, tenure, contact with the committee, awareness of governance procedures, or experience with particular programs.

**Table 3. School Committee Oversight Items with the Widest Response Dispersion (n = 18)**

| No. | School Committee Role Item                                                                     | Mean | SD   |
|-----|------------------------------------------------------------------------------------------------|------|------|
| 25  | The school committee monitors the quality of educational services provided by the school.      | 3.11 | 1.05 |
| 29  | The school committee submits recommendations based on the results of supervision.              | 3.22 | 1.03 |
| 23  | The school committee supervises the use of the school budget.                                  | 3.06 | 1.03 |
| 24  | The school committee obtains adequate information regarding the management of school finances. | 3.06 | 1.03 |
| 21  | The school committee monitors the implementation of the school program on a regular basis.     | 3.17 | 1.01 |

This distinction is important because formal participation can differ from meaningful influence. Gerrard and Savage (2022) show that parent-citizen participation in school governance can be structured in ways that differentiate forms of contribution, while Yu and Shay (2022) document how representation can become tokenistic when participants lack genuine influence. These studies do not establish what occurred in the present school, but they provide theoretical reasons to investigate whether staff have equal visibility of committee activities. A suitable follow-up would combine response-frequency reporting with interviews of committee members, school leaders, teachers, and education personnel, supported by document analysis of meeting minutes, oversight reports, budget-related communications, and program-monitoring records. Such triangulation could distinguish a communication problem from differences in respondent role or from genuinely inconsistent governance practice.

### **Interpretation, Limitations, and Transferability**

The most important contribution of the findings is diagnostic rather than generalizable. First, internal Process and Product perceptions were positively aligned in the available respondent group. Second, the Context proxy did not show a statistically reliable relationship with either internal measure. Third, the greatest item-level dispersion was concentrated in committee-oversight statements. This pattern suggests a useful local question: are internal leadership and quality-culture routines more visible and consistently experienced by staff than external oversight functions? The present data raise that question but cannot answer it causally. The finding should therefore motivate verification and qualitative inquiry rather than a definitive judgment about governance quality.

Several limitations constrain interpretation. The sample of 18 respondents is small, comes from a single school, and was not documented as probabilistically selected. Precision is limited, as shown by the wide confidence intervals around the correlations. All three constructs were measured through self-report by the same respondent group, creating possible common-method and social-desirability bias. Eight respondents could not be matched to the available official teacher roster, and their status could not be verified; consequently, the 18 respondents cannot be treated as a confirmed homogeneous teacher population. The available record also lacks instrument provenance and psychometric evidence, preventing claims about construct validity or score reliability. In addition, raw respondent-level data were unavailable during revision, so Pearson assumptions could not be independently rechecked. Detailed ethics and data-governance documentation was likewise unavailable. These limitations collectively mean that the results are exploratory and should not be generalized to schools in Ambon, Maluku, or Indonesia as a whole.

Transferability is therefore best framed as methodological and conceptual. Indonesian school leadership research itself emphasizes the importance of context in interpreting leadership practice (Lumban Gaol, 2023), while reviews of distributed leadership show substantial variation in concepts, purposes, and implementation across settings (Mifsud, 2024). Similar

schools may find value in the diagnostic approach of examining whether committee oversight is visible to staff, whether internal leadership and culture perceptions move together, and whether different respondent groups interpret governance routines consistently. However, any transfer requires local validation of instruments, verified sampling frames, explicit ethics procedures, and triangulation with administrative and qualitative evidence. A future multi-school study could test whether the pattern persists across schools, use validated scales and multi-informant data, examine measurement invariance where appropriate, and distinguish governance structures from individual perceptions of those structures.

The sensitivity of the initial  $n = 6$  interpretation also provides a methodological lesson. Very small samples can produce apparently coherent item patterns that change substantially when only a modest number of observations are added. The expanded  $n = 18$  dataset did not reproduce the earlier impression of uniformly weak committee oversight; instead, it showed relatively high aggregate Context attainment together with broader response dispersion. This reinforces the value of reporting uncertainty, avoiding categorical conclusions from sparse data, and separating descriptive patterns from explanatory narratives. It also supports the broader evaluation principle that decisions should be based on converging evidence rather than on a single indicator or a preliminary subset (Rocha et al., 2022).

## CONCLUSION

This single-school exploratory evaluation found high descriptive attainment for the three study-specific CIPP-informed indicators, with Context at 81.6%, Process at 89.7%, and Product at 88.9% of their respective maximum possible aggregate scores. Process and Product were positively associated ( $r = 0.551$ ,  $p = 0.018$ ), whereas Context was not significantly associated with either internal measure, and the widest item-level response dispersion occurred in school-committee oversight statements. These results suggest that perceptions of adaptive leadership and school quality culture were more closely aligned with each other than with perceptions of committee oversight in this respondent group, but they do not establish causal relationships, governance imbalance, or ineffective committee functioning. For local quality improvement, the school should verify staff records, document and communicate committee oversight activities more consistently, and strengthen future evaluation procedures through validated instruments, explicit ethics and data-governance documentation, response-frequency reporting, and qualitative triangulation. Future research should extend the analysis to larger multi-school samples and multiple data sources before drawing broader policy or population-level conclusions.

## REFERENCES

- Alajmi, M. (2022). School principals' experiences of autonomy and accountability: Outcomes of the School Education Quality Improvement Project in Kuwait. *International Journal of Educational Management*, 36(4), 606-617. <https://doi.org/10.1108/IJEM-09-2021-0366>
- Amels, J., Krüger, M. L., Suhre, C. J. M., & van Veen, K. (2021). The relationship between primary school leaders' utilization of distributed leadership and teachers' capacity to change. *Educational Management Administration & Leadership*, 49(5), 732-749. <https://doi.org/10.1177/1741143220915921>
- Bendermacher, G. W. G., Dolmans, D. H. J. M., de Grave, W. S., Wolfhagen, I. H. A. P., & Oude Egbrink, M. G. A. (2021). Advancing quality culture in health professions education: Experiences and perspectives of educational leaders. *Advances in Health Sciences Education*, 26(2), 467-487. <https://doi.org/10.1007/s10459-020-09996-5>

- Bush, T. (2021). School leadership and culture: Societal and organisational perspectives. *Educational Management Administration & Leadership*, 49(2), 211-213. <https://doi.org/10.1177/1741143220983063>
- Cahyadi, A., Hendryadi, Widyastuti, S., & Suryani. (2022). COVID-19, emergency remote teaching evaluation: The case of Indonesia. *Education and Information Technologies*, 27, 2165-2179. <https://doi.org/10.1007/s10639-021-10680-3>
- Demirhan, G. (2023). School culture types as predictors of school leadership capacity. *International Journal of Contemporary Educational Research*, 10(3), 715-723. <https://doi.org/10.52380/ijcer.2023.10.3.535>
- Dizon, A. G. (2023). Historical development of CIPP as a curriculum evaluation model. *History of Education*, 52(1), 109-128. <https://doi.org/10.1080/0046760X.2022.2098390>
- Dwangu, A. M., & Mahlangu, V. P. (2021). Accountability in the financial management practices of school principals. *International Journal of Educational Management*, 35(7), 1504-1524. <https://doi.org/10.1108/IJEM-06-2021-0243>
- Gerrard, J., & Savage, G. C. (2022). The governing parent-citizen: Dividing and valorising parent labour through school governance. *Journal of Education Policy*, 37(5), 744-761. <https://doi.org/10.1080/02680939.2021.1877357>
- Heifetz, R. A., & Linsky, M. (2002). *Leadership on the line: Staying alive through the dangers of leading*. Harvard Business School Press.
- Hutt, E., & Polikoff, M. S. (2020). Toward a framework for public accountability in education reform. *Educational Researcher*, 49(7), 503-511. <https://doi.org/10.3102/0013189X20931246>
- Kasmairi, K., Hamzah, S., & Winarto, H. (2023). CIPP evaluation model: Online in-service teachers training program conducted at English Language Education Study Program of Bengkulu University. *English Franca: Academic Journal of English Language and Education*, 7(1), 105-118. <https://doi.org/10.29240/ef.v7i1.5871>
- Kirk, G., & Barblett, L. (2022). Implementing the National Quality Standard in schools: Leadership that motivates improvement initiatives through psychological ownership. *The Australian Educational Researcher*, 49, 367-385. <https://doi.org/10.1007/s13384-021-00446-8>
- Kumalasari, B., & Idawati, L. (2023). Evaluasi program pendidikan karakter di SD Athalia dengan model CIPP (Context, Input, Process, Product). *Jurnal Akuntabilitas Manajemen Pendidikan*, 11(2), 60-72. <https://doi.org/10.21831/jamp.v11i2.65007>
- Leithwood, K., Sun, J., & Schumacker, R. (2020). How school leadership influences student learning: A test of The Four Paths Model. *Educational Administration Quarterly*, 56(4), 570-599. <https://doi.org/10.1177/0013161X19878772>
- Liu, Y., Bellibaş, M. Ş., & Gümüş, S. (2021). The effect of instructional leadership and distributed leadership on teacher self-efficacy and job satisfaction: Mediating roles of supportive school culture and teacher collaboration. *Educational Management Administration & Leadership*, 49(3), 430-453. <https://doi.org/10.1177/1741143220910438>
- Lumban Gaol, N. T. (2023). School leadership in Indonesia: A systematic literature review. *Educational Management Administration & Leadership*, 51(4), 831-848. <https://doi.org/10.1177/17411432211010811>
- Luschei, T. F., & Jeong, D. W. (2021). School governance and student achievement: Cross-national evidence from the 2015 PISA. *Educational Administration Quarterly*, 57(3), 331-371. <https://doi.org/10.1177/0013161X20936346>

- Matete, R. E. (2022). Forms of decentralization and their implications for educational accountability in Tanzania. *Heliyon*, 8(5), e09436. <https://doi.org/10.1016/j.heliyon.2022.e09436>
- Mifsud, D. (2024). A systematic review of school distributed leadership: Exploring research purposes, concepts and approaches in the field between 2010 and 2022. *Journal of Educational Administration and History*, 56(2), 154-179. <https://doi.org/10.1080/00220620.2022.2158181>
- Nuraya, N., Adiman, A., Herawati, E. S. B., Warlizasusi, J., & Aliyyah, R. R. (2021). Implementation of the roles and functions of the school committee in improving the quality of education services in the City of Cirebon. *Tadbir: Jurnal Studi Manajemen Pendidikan*, 5(1), 81-100. <https://doi.org/10.29240/jsmp.v5i1.2113>
- Ordofa, B., & Asgedom, A. (2022). School accountability and its relationship with learning outcomes: A systematic literature review. *Social Sciences & Humanities Open*, 6(1), 100358. <https://doi.org/10.1016/j.ssaho.2022.100358>
- Regulation of the Minister of Education and Culture of the Republic of Indonesia Number 75 of 2016 concerning School Committees.
- Rocha, A. C., Silva, M., & Duarte, C. (2022). How is sexuality education for adolescents evaluated? A systematic review based on the Context, Input, Process and Product (CIPP) model. *Sex Education*, 22(2), 198-216. <https://doi.org/10.1080/14681811.2021.1908984>
- Sallis, E. (2002). *Total quality management in education (3rd ed.)*. Kogan Page.
- Sari, I. K., Kasmini, L., Rosdiana, R., & Manurung, M. M. H. (2023). The online evaluation of the teacher certification program using the CIPP model. *Jurnal Penelitian dan Evaluasi Pendidikan*, 27(2), 164-176. <https://doi.org/10.21831/pep.v27i2.57914>
- Stufflebeam, D. L. (2003). The CIPP model for evaluation. In T. Kellaghan & D. L. Stufflebeam (Eds.), *International handbook of educational evaluation* (pp. 31-62). Springer.
- Sumiati, S., Kurniady, D. A., Komariah, A., Suryana, A., Dadi, D., & Hafidh, Z. (2024). Linking principal adaptive leadership to teacher performance: The mediating effect of collaborative school culture. *Journal of Social Studies Education Research*, 15(4), 17-41.
- Supriadi, D., Usman, H., Jabar, C., & Widyastuti, I. (2021). Good school governance: An approach to principal's decision-making quality in Indonesian vocational school. *Research in Educational Administration & Leadership*, 6(4), 796-831. <https://doi.org/10.30828/real/2021.4.2>
- Tan, C. Y., Dimmock, C., & Walker, A. (2024). How school leadership practices relate to student outcomes: Insights from a three-level meta-analysis. *Educational Management Administration & Leadership*, 52(1), 6-27. <https://doi.org/10.1177/17411432211061445>
- Tolo, A., Lillejord, S., Flórez Petour, M. T., & Hopfenbeck, T. N. (2020). Intelligent accountability in schools: A study of how school leaders work with the implementation of assessment for learning. *Journal of Educational Change*, 21, 59-82. <https://doi.org/10.1007/s10833-019-09359-x>
- Yu, K., & Shay, C. (2022). Tokenism and barriers to genuine learner participation in school governance in one progressive South African girls' high school. *South African Journal of Education*, 42(4), Article 2115. <https://doi.org/10.15700/saje.v42n4a2115>