Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Local Education Governance and Principal Curriculum Management Capacity Across Districts and Schools in Indonesia

Muhammad Soleh[1]

muhamadsholeh@unesa.ac.id

Universitas Negeri Surabaya

Hafis Muaddab[2]

State University of Malang

Khudrotun Nafisah

Universitas Darul Ulum

Ahmad Abdullah Zawawi

Universitas Negeri Surabaya

Research Article

Keywords: local education governance, decentralisation, curriculum reform, principal capacity, professional learning, cascade training, school leadership

Posted Date: November 19th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11285303/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

Indonesia's education system operates under multi-level governance, pairing national curriculum standards with decentralised district management. This study asks whether local governments can translate national intent into school practice. It analyzes the 2023 National Assessment for 38,579 primary and lower-secondary schools in 488 districts. Principals' curriculum management capacity is the lowest-scoring of 39 principal-level indicators: more than 85 percent of principals score below the adequacy threshold, and no district reaches an adequate average at lower-secondary level. School resources do not explain this deficit, since sixteen structural conditions account for only 1.4 to 3.1 percent of the variation. Districts explain little either. Multilevel models show that districts account for 2 to 4 percent of the variation, and none of the 39 indicators has a between-district share above 8 percent. The remaining 96 to 98 percent lies between schools within the same district. This pattern is consistent with a training cascade that ends at each school separately, although it is not specific to curriculum management. Accounts from 52 principals in one well-provisioned regency point the same way: principals can state what the reform requires but cannot translate it into instructional guidance, and they name training, not resources, as the constraint. The near-absence of district-level differentiation indicates that local governance is not yet functioning as a professional learning system. Districts could act on this using instruments they already hold: school-level assessment data, supervisors, and principal working groups.

1 Introduction

Indonesians have a saying for curriculum reform: *ganti menteri, ganti kurikulum*, a new minister means a new curriculum [45]. The country has had at least ten national curricula since 1947 [27], and each has been handed to the same people to carry out: the principals of more than 437,000 schools, spread across 552 subnational governments [52]. This paper asks what those principals can do with a curriculum, and whether the local government they answer to makes any difference.

The question is not peculiar to Indonesia. Reform rarely reaches classrooms in the form its designers intend. OECD [46] found that hundreds of reforms introduced across member countries between 2008 and 2014 produced relatively few changes in classroom practice, and Sahlberg [55] notes that the effects of standardisation and external accountability on learning have been uneven and fragile. The common diagnosis is that reform has been treated as a technical matter when it depends on professional and relational conditions [17, 56]: educators engage with change they experience as inclusive and comply superficially with change they experience as imposed [24], and trust within and between schools and districts shapes which of the two occurs [39, 41].

Indonesia makes the problem visible at unusual scale. Its 3.3 million teachers and 53 million students are distributed across two ministries [52], the curriculum architecture is centralised, and each reform arrives before the previous one has settled [54]. It is a system that has reformed continuously without, in

many respects, reforming deeply. Principals — the people on whom curriculum enactment depends — are repeatedly asked to lead change for which they have not been prepared.

Reform in Indonesia is also layered onto a fiscal architecture that was not designed to carry it. Decentralisation arrived abruptly after 1998, driven by fiscal crisis and regional political pressure rather than by coherent administrative design, and its consequences for educational attainment have been uneven [42, 69]. Districts hold budgetary authority over schooling but receive most of their revenue as block transfers whose largest share is committed to personnel costs before any discretionary decision is made [38, 68]. The administrative unit through which curriculum reform must pass is therefore one that holds formal authority over spending and limited practical room to redirect it.

The current reform has been assessed favourably on precisely the terms this literature recommends. Reviewing Merdeka Belajar for the OECD, Radinger and Hague [52] describe a reform that narrows the curriculum towards foundational competencies, devolves instructional decisions to teachers, invests in school leadership, and was adopted voluntarily rather than imposed — features they identify as having been critical to reform success in other systems. Their assessment is also explicitly provisional: the reform's true test, they conclude, lies in implementation. Two observations in that review frame the present study. The first is that national standards for Indonesian principals emphasise managerial competence over pedagogical leadership. The second is their statement that further data would be needed to establish the quality of principal–teacher interaction and how leadership is actually distributed across schools. The data analysed here address that gap directly, and at the scale of the system rather than the sample.

This paper examines what happens at the end of that chain. Using national assessment data for 38,579 schools alongside qualitative accounts from 52 principals, we ask: where does the deficit in curriculum management capacity sit, and what does its distribution reveal about how local education governance works — or fails to work — in a large, decentralised system with a centralised curriculum? The question matters beyond Indonesia. Curriculum management — the principal's capacity to develop and manage the school's curriculum in ways that improve student learning — is precisely the kind of leadership practice that reform depends on. If it does not follow resources, and is not differentiated at district level, then the dominant models of reform support are targeting the wrong things. The Indonesian data make this visible at national scale; the mechanism they point to is familiar from research conducted in very different contexts [44, 48, 56].

We structure the paper as follows. Section 2 sets out the theoretical framing and data. Section 3 presents findings. Section 4 discusses what the results mean for local education governance and for how educational change is supported. Section 5 concludes with implications for districts and for national assessment.

2 Theoretical Framing and Methods

2.1 Three accounts of the last mile

Research on educational change has long recognised a gap between policy intent and enacted practice. Spillane et al. [64] showed that reform implementation is fundamentally a cognitive process: what principals and teachers do with a reform depends on how they interpret it, shaped by prior knowledge, local conditions, and the meaning-making that happens in schools. Cuban [11] documented how reforms are absorbed into existing practice rather than displacing it. Priestley et al. [51] argue that curriculum is not a text schools receive and deliver but a set of meanings constructed at every level of a system, with the school as the critical site of reconstruction. Whether principals can lead this reconstruction is a question that turns less on what their schools have than on what they have had the opportunity to learn.

A growing body of scholarship locates reform failure not in inadequate design but in the absence of relational and professional conditions that make change meaningful. Sahlberg [56] argues that trust operates simultaneously as a precondition for change, a mechanism through which change is enacted, and an outcome reflecting the quality of reform processes. Niedlich et al. [44], in a comprehensive literature review, show that trust mediates between policy and practice, shaping whether educators perceive reform as legitimate and invest effort in genuine enactment. Ehren et al. [15] argue that effective accountability systems must be grounded in professional trust rather than control. We draw on this literature to interpret our findings; we do not measure trust directly.

Three explanations generate testable predictions about where a deficit in principal capacity should sit.

The resource explanation holds that what schools can do follows from what schools have — the logic behind profiling systems, equity funding, and targeted support [10, 23]. The relationship between spending and outcomes has long been contested [22, 29], and Indonesia has tested its strong form directly: de Ree et al. [13] found that an unconditional doubling of certified teachers' salaries produced no measurable improvement in student learning. If correct, curriculum management capacity should vary with structural conditions and cluster geographically. In a decentralised system the second half of this prediction is stronger still, since fiscal devolution is widely argued to widen disparities between subnational units [2, 16].

The governance explanation locates capacity in local government. In Indonesia, districts and cities hold authority over primary and lower-secondary schooling [53]: they appoint and rotate principals, employ the school supervisors (pengawas) who are the system's designated source of instructional support, allocate the discretionary share of the education budget, and host the principal working groups (KKKS and MKKS) through which peer learning is meant to occur. How districts use the first of these powers is itself in question. A systematic review of sixteen studies of Indonesian school leadership published between 2004 and 2019 identifies a lack of capacity to lead and manage schools as the most frequently reported problem and inappropriate principal selection as another [40]; in the study that review draws on most for the latter, principals in four regions were appointed chiefly through personal connections with district officers, with little regard to training or qualification [66]. There is evidence that districts do differ.

Comparing three Indonesian districts, Arif et al. [3] show that some local governments pursue policies aimed at the quality of learning and others do not, and that the difference follows from local politics: where the district head's discretion over bureaucratic appointments is constrained, education offices develop the capacity to design such policies themselves; where it is not, quality-oriented policy depends on the personal commitment of the district head and is exposed to reversal at the next election. If districts differ in how well they use these instruments to translate national curriculum policy, principals' capacity should cluster by district: the between-district share of variance should be substantial, and district-level characteristics such as education budget share, minimum service standard status and fiscal capacity should explain it. If, instead, districts act mainly as relay points that pass a centrally designed socialisation on to schools, district membership should explain little, and capacity should vary school by school.

The professional learning explanation holds that capacity follows from what staff have had the opportunity to learn together [25, 74]. In Indonesia, this operates through a cascade: training delivered centrally to officials who pass it on without adequate support, losing depth at every step [26, 54]. The cascade lacks the features the professional development literature identifies as necessary: differentiation, follow-up, and connection to classroom practice [12, 14]. If correct, structural conditions should explain little, and variation should lie between schools within districts rather than between them — because the cascade ends at each school individually.

2.2 Data and analytic approach

We used three sources. The first is the 2023 Asesmen Nasional principal-level public report, covering 24,789 primary and 13,790 lower-secondary schools across 488 districts. The outcome variable (MIP) measures the principal's capability to develop and manage the curriculum to improve student learning — a composite of principal and teacher responses, scaled 0 to 100. Conceptually, the indicator corresponds most closely to what Hallinger and Murphy [20] term managing the instructional program, the dimension of instructional leadership that covers coordinating the curriculum, supervising instruction and monitoring student progress, as distinct from defining the school's mission or promoting a positive learning climate. The structural predictor block spans sixteen variables: teacher qualification and certification, school size, pupil-teacher ratio, computers, libraries, poverty assistance receipt, socio-economic index, electricity, internet access, urban-rural location, special-region status, public-private status, curriculum version, and macro-region. Table 1 reports their distribution. A supplementary block of nine survey composites from the same instrument tested the ceiling of within-instrument explanation.

Table 1
Characteristics of the analytic samples.

Characteristic	Primary (n = 24,789)	Lower-secondary (n = 13,790)
Teachers with at least a bachelor's degree (%)	96.3	97.3
Certified teachers (%)	33.3	28.5
Students (mean)	168.4	260.9
Pupil-teacher ratio (mean)	15.4	15.0
Computers owned (mean)	0.27	0.56
Students receiving PIP assistance (%)	47.3	47.1
School socio-economic index (mean)	70.9	71.2
Internet available (%)	70.9	67.3
Public schools (%)	83.7	54.8
Urban location (%)	30.4	37.8
Using Kurikulum Merdeka in 2023 (%)	76.1	71.5

The second source is the 2025 Neraca Pendidikan Daerah for Gresik Regency, East Java, and Indonesia [31, 32, 33], allowing us to characterise our qualitative site relative to national benchmarks. The third is 780 open-ended responses from 52 principals in Gresik, analysed using reflexive thematic analysis [6, 7] with two independent coders. Participation was voluntary and responses were anonymised.

We estimated four multilevel models at each school level. Model 0 (null model) yields the intraclass correlation. Model 1 regresses the outcome on the structural block. Model 1b adds district fixed effects — an upper bound on what structural conditions plus all unmeasured district factors can explain. Model 2 adds the nine survey composites; Model 2a removes school vision and mission given its conceptual proximity to the outcome. With samples of this size, statistical significance is uninformative; we report standardised coefficients, explained variance, and effect sizes, interpreting magnitude rather than significance. Listwise deletion left 22,982 primary schools (92.7 percent) and 10,984 lower-secondary schools (79.7 percent); excluded schools scored lower on the outcome, so estimates understate the deficit.

3 Findings

3.1 The weakest capacity

Curriculum management scores are low and, remarkably, low almost everywhere. The primary mean is 48.71 (SD = 15.43, median = 43.79, IQR = 39.03–55.01); the lower-secondary mean is 47.67 (SD = 15.31). Sixty-four and 66 percent of principals respectively score below 50; 85 and 86 percent score below 60; only 11 and 10 percent reach 70. The distribution is positively skewed and uneven: most schools sit between 30 and 60, and a separate cluster of about 8 to 9 percent of schools scores between 85 and 90, with few schools in between (Fig. 1). Among 39 principal-level composite indicators, curriculum management ranks lowest at both school levels (Table 2) — not a domain where some schools struggle while others thrive, but the bottom of a system-wide distribution.

Table 2
Six lowest-scoring principal-level indicators, National Assessment 2023 (0 to 100).

Rank	Primary	Mean	Lower-secondary	Mean
1	Curriculum management	48.71	Curriculum management	47.67
2	Sexual violence policy	52.58	Knowledge of gifted students	53.35
3	Open classroom climate	52.99	Learning about learning	53.39
4	Gender equality attitudes	54.15	Understanding of physical punishment	54.69
5	Knowledge of gifted students	54.37	Open classroom climate	54.73
6	Learning about learning	54.47	Gender equality attitudes	54.85

At district level the pattern is stronger still. Across 478 districts with at least ten sampled primary schools, means run from 39.67 to 60.07; at lower-secondary level, 38.20 to 59.68. At primary level 99.8 percent of districts fall below 60; at lower-secondary level all of them do (Fig. 2). No district at lower-secondary level, and only one at primary level, reaches an adequate average. This is not a finding about disadvantaged schools in neglected regions. It is a finding about the system.

The variation that exists is strikingly flat across the structural predictors equity policy relies on (Table 3, Fig. 4). The richest tenth of schools scores only 3.80 points above the poorest tenth at primary level — about a quarter of one within-group standard deviation. Urban schools outscore rural by 2.24 points. Schools with internet outscore those without by 1.54. These differences are negligible against an outcome standard deviation of 15.

Table 3
Curriculum management by school subgroup and socio-economic decile.

Comparison	Primary	Lower-secondary
Urban / rural	50.26 / 48.02	49.08 / 46.81
Public / private	48.93 / 47.57	49.22 / 45.78
Internet available / not	49.32 / 47.78	—
Special region / not	46.88 / 48.79	45.03 / 47.77
Kurikulum Merdeka / Kurikulum 2013	49.16 / 47.26	48.58 / 45.38
Socio-economic decile 1 (lowest)	46.82	45.36
Socio-economic decile 5	48.89	47.06
Socio-economic decile 10 (highest)	50.62	50.86

3.2 What schools have

Sixteen structural conditions explain 1.4 percent of the variation at primary level and 3.1 percent at lower-secondary level (Table 4, Fig. 3). Individual coefficients confirm this (Table 5): urban location is worth 1.09 points; school wealth, 0.72 per standard deviation; teacher certification, 0.30; computer ownership, 0.12. Against an outcome standard deviation of about 15, all are negligible. What schools have does not predict what their principals can do with a curriculum.

Table 4
Variance explained, by model.

Model	Predictors	Primary	Lower-secondary
1	Structural block (16 variables)	.014	.031
1b	Structural block + district fixed effects	.063	.094
2	Structural block + 9 survey composites	.601	.587
2a	Model 2 without vision and mission	.477	.466

Table 5
Model 1 coefficients, in outcome points.

Predictor	Primary	Lower-secondary
Private (against public)	−0.51	−1.89
Urban (against rural)	1.09	0.71
Special region	0.71	0.32
No internet	−0.41	−0.13
Kurikulum Merdeka (against 2013)	0.98	1.23
Central Indonesia (against western)	1.17	0.64
Eastern Indonesia (against western)	2.69	1.58
Teachers with bachelor's degree	0.05	0.17
Certified teachers	0.30	0.32
Number of students	0.92	1.27
Pupil-teacher ratio	0.14	−0.20
Learning groups	−0.05	0.10
Computers owned	0.12	0.23
Libraries	0.26	−0.02
Students receiving PIP	0.54	0.02
School socio-economic index	0.72	0.99

Note. Continuous predictors standardised. Analytic n = 22,982 and 10,984. Significance tests omitted deliberately: at this sample size most coefficients pass conventional thresholds while remaining practically negligible.

Adding district fixed effects in Model 1b raises explained variance to only 6.3 and 9.4 percent — meaning that even after absorbing everything distinguishing one district from another, about 90 percent of variation remains unaccounted for. The multilevel decomposition makes the asymmetry clear: the structural block explains 19.3 percent of between-district variance at primary level but only 0.6 percent of within-district variance; 12.4 and 2.7 percent at lower-secondary. School conditions describe the small part of variation that follows geography. They say almost nothing about the large part that does not.

3.3 Where schools are

The intraclass correlation is .036 at primary level and .023 at lower-secondary level: between 96 and 98 percent of the variation lies between schools within the same district. The three accounts make different predictions about this quantity. A resource account predicts that capacity follows structural conditions,

which cluster geographically. A governance account predicts clustering by district. A cascade account predicts neither, because the chain ends at each school individually. The observed pattern is most consistent with the third.

Two cautions apply. First, measurement error in a survey-based composite accumulates at the school level and inflates the within-district share. The public release contains composite scores only, without item-level data, so the reliability of the indicator cannot be estimated and the intraclass correlations reported here are uncorrected. Second, a low intraclass correlation is informative only if it is distinctive, and it is not. Estimating the same quantity for all 39 principal-level indicators (one-way analysis-of-variance estimator) gives between-district shares from .007 to .079 at primary level (median .038) and from .001 to .066 at lower-secondary level (median .033). Curriculum management, at .032 and .021 on this estimator, sits in the middle of that range: 17th lowest of 39 at primary level and 9th lowest at lower-secondary level. What distinguishes curriculum management is therefore its level, the lowest mean of the 39 indicators, and not its distribution. The within-district pattern is a property of every principal capacity the assessment records. This broadens the governance finding, since districts do not differentiate themselves on any measured capacity, but it weakens the inference to a curriculum-specific mechanism: a cascade that ends at each school is one explanation of the pattern, and a general absence of district influence on principals, or measurement error common to the instrument, are others that these data cannot rule out. The same structural model applied to the professional learning indicator yields 3.2 and 6.3 percent explained variance, which suggests the insensitivity to school conditions is not unique to curriculum management.

For local governance, the implication is specific. Districts differ in budget, staffing and fiscal capacity, yet they do not differ much in what their principals can do with a curriculum. Whatever districts currently do in this domain, they do it in much the same way, and it does not reach schools differentially. We could not test district-level predictors such as education budget share, minimum service standard status or fiscal capacity directly, because the public release anonymises district codes and cannot be linked to fiscal data. The limitation is bounded, however: no district characteristic, measured or unmeasured, can account for more than the between-district share of 2 to 4 percent, and the district fixed effects in Model 1b already absorb all of them.

3.4 What principals say

Gresik Regency outperforms the national average on teacher certification (73.0 vs 63.5 percent at primary), internet access (98.7 vs 87.3 percent), top accreditation grade (55.6 vs 21.7 percent of primary schools), literacy, numeracy, and dropout (Table 6). If capacity followed provision, it should show more of it. The 52 principals in this regency report the same deficit the national data document.

Table 6
Gresik Regency, East Java and Indonesia compared, 2025.

Indicator	Gresik	East Java	Indonesia
Certified teachers, primary (%)	73.0	73.1	63.5
Certified teachers, lower-secondary (%)	63.3	64.4	61.7
Internet available, primary (%)	98.7	96.2	87.3
Internet available, lower-secondary (%)	99.2	95.4	92.2
Accreditation grade A, primary (%)	55.6	20.1	21.7
Accreditation grade A, lower-secondary (%)	59.1	41.4	32.1
Schools rated good on literacy, primary (%)	68.3	57.6	42.4
Schools rated good on numeracy, primary (%)	51.3	46.8	32.7
Dropouts per 10,000, primary	4.6	5.1	13.7
Net enrolment rate, lower-secondary (%)	81.1	79.4	76.0
Education share of local budget (%)	30.2	29.7	—
Minimum service standard status	Tuntas Madya	Tuntas Pratama	—

Source. Neraca Pendidikan Daerah 2025 [31, 32, 33]. Dropout rates calculated from reported counts and enrolments.

The district's own documents show the same pattern from the provider's side. The 2023 budget ceiling of the Gresik education office was Rp 951.9 billion [4]. Of this, 50.7 percent was allocated to the programme supporting local government affairs, almost all of it under financial administration, and 44.6 percent to education management by school level. The programme for educators and education personnel received 4.6 percent, and the curriculum development programme received Rp 638 million, or 0.07 percent, all of it for setting local-content curricula. No budget line is identifiable as support for principals' curriculum or instructional leadership. The same report, prepared by the regency's planning agency from the 2023 National Assessment results for lower-secondary schools (273 schools, 207 responding principals), singles out instructional leadership as an area needing improvement: it describes it as limited, not yet driving planning, practice and assessment oriented to student learning, and not yet developing the programmes, incentives and resources that support teachers in reflecting on and improving their teaching. The regency has therefore identified the deficit in its own planning documents; what the budget does not yet show is a channel for addressing it.

The accounts refer to Deep Learning (Pembelajaran Mendalam), the pedagogical approach introduced by the ministry in 2025 [34], and therefore to a later reform step than the 2023 assessment data. We

treat them as illustrating the mechanism by which reform reaches principals, not as a measurement of the same moment.

On the closed items, training and socialisation was the weakest capacity sub-dimension (57.7), and the weakest single item concerned whether teachers had received comparable preparation (54.5). Three themes emerged from the open-ended responses.

First, the cascade fails at transmission. Principals attended one or two socialisation sessions, then were expected to pass what they absorbed to their teachers:

The training exists, but not all the teachers at our school get the same chance. Usually the principal is invited to the socialisation first, and then it gets passed along to the teachers. So, how much the teachers understand depends entirely on how much we managed to take in during that socialisation. (P3)

The socialisation on Deep Learning did happen, but the time was very short. Straight afterwards we were asked to pass this understanding on to the teachers. Honestly, we ourselves are sometimes still unsure and do not fully understand how to apply it in practice. (P1)

Second, conceptual vocabulary has arrived but operational capacity has not — what Shaked [60] calls a failure of instructional leadership clarity:

Conceptually we do understand it, learning has to be mindful, meaningful and joyful. But the moment someone asks, so in a Year 4 mathematics lesson what does it actually look like, that is where we start struggling to give a really concrete example. (P2)

A similar pattern has been reported in rural Bangladesh, where principals rated themselves highest on defining the school mission and lowest on observing and improving instruction [61].

Third, commitment held while capability lagged. Orientations towards the reform scored 78.0, understanding 65.4, school capacity 65.5, with training at 57.7 within it — the pattern Weiner [74] predicts when commitment can be raised quickly by messaging while capability requires sustained professional learning. Principals did not reject the reform. They described it as a good idea they had not been equipped to enact:

What we need now is not just being gathered together to listen to a socialisation about what Deep Learning is. We need to be guided, how to turn this good concept into concrete steps when we are actually teaching in class. (P4)

4 Discussion: governance without a learning system

Why do reforms that are carefully designed and sincerely pursued so often fail to change what happens in schools? The Indonesian data offer a view at national scale, because the variance structure bears on

competing accounts and does more than document a deficit.

The resource explanation has a long pedigree and a persistent institutional logic. Leithwood et al. [37], in a foundational framework for research on large-scale reform, distinguished between reforms that change the structural conditions of schooling and those that build the professional capacity of the people within it — and argued that the latter is the more reliable lever for sustained improvement. School profiling systems, equity funding models, and targeted improvement programmes are all built on the former assumption: that what schools can do follows from what schools have [10, 23]. The Indonesian data do not support it for curriculum management. Sixteen input variables account for one to three percent of a capacity whose variation is 96 to 98 percent within districts. The deficit does not follow the resource gradient and is not concentrated where equity instruments are pointed.

This result also speaks to a prediction from outside the school effectiveness literature. Fiscal decentralisation is widely held to widen disparities between subnational units [2, 16], and Indonesian transfers do vary across districts in both volume and discretion [19, 38]. Intraclass correlations of .036 and .023 are not what that literature would lead one to expect. Whatever curriculum management capacity tracks, it is not the dimension along which decentralised fiscal capacity varies. One reading is that the composition of educational spending matters more here than its volume or its distribution. Sumiyana and Effendi [68] describe an Indonesian budgeting system in which block transfers are largely absorbed by personnel costs and remaining discretion is directed towards physical infrastructure rather than capability development. A system organised this way can fund a socialisation event; it has no established channel through which to fund sustained, school-based professional learning. The Gresik budget fits this description: half of the education office's ceiling sits in administrative support, and curriculum development receives 0.07 percent [4]. A constraint that is uniform would produce a deficit that is uniform.

This leaves a puzzle. Case-study evidence shows that Indonesian districts vary in bureaucratic capacity and in the education policies they adopt [3], yet district membership accounts for only 2 to 4 percent of the variation in principals' curriculum management. Three readings are compatible with both findings. First, the quality-oriented policies districts adopt rarely target what principals do with a curriculum; those documented by Arif et al. concern teacher attendance, teacher training, parental involvement and grade promotion. Second, such policies are fragile. In one of their three districts a programme sustained for a decade was discontinued shortly after a change of district head, a local counterpart to the national pattern of ganti menteri, ganti kurikulum. Third, districts with the bureaucratic capacity to generate their own instructional policy may simply be few. Our data cannot separate these readings, but each implies the same thing: variation in local governance exists, and it is not yet reaching the capacity this reform depends on.

The professional learning explanation fares better. The variance structure — almost entirely within districts — is what a cascade model predicts. A chain of transmission that terminates at each school separately, regardless of geography or governance context, produces the pattern we observe. But the

cascade is not simply an inefficient delivery mechanism. Tikkanen et al. [71], studying large-scale curriculum reform in Finland, found that reform succeeds when it builds collective responsibility, mutual trust, and inclusive membership — creating spaces in which educators work together toward common goals. The Indonesian cascade is the structural inverse: it delivers reform concepts to each school individually, without the collegial dialogue, follow-up, or practice-connected engagement that would convert those concepts into shared professional capacity. Owen and Wong [48], examining a parallel case in Timor-Leste, documented this failure mode: a cascade in which school leaders received initial training and were then expected to pass it to teachers without adequate support, producing transmission that was thin, inconsistent, and unable to reach classroom practice. Their solution — combining workshop series, peer learning groups, and on-site mentoring — points to the infrastructure the Indonesian data suggest is absent. This is consistent with what is known about how leadership reaches classrooms. Thoonen et al. [70] found that teaching practices improve through teachers' engagement in professional learning, which leadership and organisational conditions shape indirectly. Neumerski [43], reviewing research on principal, teacher and coach instructional leadership, concluded that the field knows far more about what such leaders are expected to do than about how they do it. A socialisation session transmits the first kind of knowledge and none of the second. Evidence from Indonesian teachers points to the same missing element: in a survey of 2,531 in-service teachers recruited through an online professional learning network, respondents valued online courses for resources and pedagogical knowledge but gave institutional support the lowest ratings [28].

This has a specific consequence for the reform's central design move. Kurikulum Merdeka loosens national prescription: teachers are released from a fixed sequence of topics and schools are invited to adapt national standards to their own context [52]. Indonesian schools already reported curriculum autonomy at or slightly above the OECD average before the reform; what the reform adds is not the freedom but the expectation that it be exercised deliberately. An earlier study of school-based management in Indonesia found the same pattern: most principals reported having autonomy over school decisions but did not use it to make significant instructional changes, and the authors called for developing district capacity to support schools [73]. The capacity to exercise that autonomy is what our outcome measures, and it is the weakest of the 39 capacities the national assessment records. Devolved curriculum authority does not become meaningful at the moment it is granted. Where the capability to use it is absent system-wide, devolution relocates the interpretive burden to the school without relieving it — an outcome the curriculum literature anticipates [47, 51] but which has rarely been observed at national scale.

The remedy Indonesia has chosen is instructive for the same reason. The Master Teacher programme (Pendidikan Guru Penggerak) trained 2,800 teachers across 58 subnational governments in its initial phase, with plans at the time of the OECD review to reach over 400,000 teachers, roughly 12 percent of the national cohort; selection is competitive and graduates are given priority for principal and superintendent posts [52]. As a strategy for cultivating a leadership cadre this is defensible. As a strategy for closing a deficit whose variance lies 96 to 98 percent between schools within the same district, it is aimed at the wrong distribution. A competitively selected minority is drawn, by construction,

from the upper tail of the distribution we observe; it cannot substitute for professional learning that reaches each school. An early evaluation of the programme likewise found that principals had not yet observed measurable school-level changes [36]. The binding constraint our data identify is not the quality of the leaders such a programme produces, but the absence of anything comparable for the schools it will not reach.

A further explanation lies in how the principal's role is defined, and in what that definition leaves out. Curriculum management is unusual among the 39 capacities the national assessment records: it is the only one whose substance is redefined by each successive reform. When teachers deliver a centrally written syllabus, the principal's curriculum task is to secure compliance with it. When teachers are expected to construct the curriculum for their own context, the principal's task becomes assuring the quality of the interpretations their teachers produce. These are different jobs calling on different capabilities, and Indonesia has alternated between them repeatedly. The 2004 competency-based curriculum positioned teachers as implementers; the 2006 school-level curriculum added a developer role; the 2013 curriculum narrowed it again to delivery of centrally prepared syllabi; Kurikulum Merdeka restored it [27]. Each of these shifts redefined what a principal was for. None of them did so explicitly. Hamaniuk et al. [21] describe a comparable dynamic in Ukrainian science education, where successive reforms accumulated as layers without replacing their predecessors, and argue that friction between those layers is one source of the implementation gap. The Indonesian principal's curriculum role, redefined four times in twenty years and never restated, can be read as such a layering.

The regulation that constitutes the office confirms the omission. The rules governing the post have themselves changed repeatedly, from Kepmendiknas 162/U/2003 through Permendiknas 13/2007 and 28/2010 to Permendikbud 6/2018, so that candidates must continually adapt to new selection procedures [62]. Permendikbudristek 40/2021 defined the principal as a teacher assigned 'to lead learning and manage the school' [30, Art. 1], made the Master Teacher certificate a condition of appointment, waivable only where certified teachers were in short supply (Arts. 2 and 4), and directed the principal's workload of management, entrepreneurship and supervision of teachers towards student-centred learning and the quality of learning outcomes (Art. 12). Even so, the word 'curriculum' does not appear in it. Permendikdasmen 7/2025, which replaced it, goes further. It specifies the competencies required of a principal in a single clause: social, personal and professional, the last of these defined as including entrepreneurial capability [35, Art. 2]. Neither instructional nor curriculum leadership appears. The eligibility conditions are administrative: academic qualification, teaching certification, rank, a performance rating, and at least two years of managerial experience (Art. 7). Across thirty-four articles, none states what a principal is expected to do with a curriculum. The same regulation removed the Master Teacher requirement, decoupling appointment from the one national programme designed to build pedagogical leadership, and shortened maximum tenure from the four four-year terms allowed in 2021 [30, Art. 8] to two (Arts. 23 and 33). The document defines the office administratively and exhaustively: how principals are selected, appointed, extended, reassigned and dismissed. What it does not define is the work. This is the documentary counterpart to what Radinger and Hague [52] observed about Indonesian principal standards, and it is consistent with our finding that the weakest measured

capacity is the one no instrument of the system tells principals how to exercise. The gap extends to the district level: Sumintono et al. [67], discussing the district's role in supporting principals' instructional leadership in Indonesia, report evidence that many district education heads and supervisors lacked an adequate understanding of the regulations they were expected to apply.

We did not measure trust, and the data cannot show whether its absence is part of the mechanism. The literature nonetheless suggests why a cascade of this kind would produce compliance without capability. Niedlich et al. [44] show that trust mediates between policy and practice, shaping whether educators perceive reform as legitimate and invest effort in enactment; Louis [39] and Hargreaves [24] show that educators engage more deeply with change they experience as inclusive than with change they experience as imposed. Ehren et al. [15] and Tolo et al. [72] argue that accountability works when leaders combine professional trust with challenge and support. Saunders [59] showed that being held accountable for reform outcomes without the support to achieve them carries an emotional cost. The Gresik accounts are consistent with this: commitment is high (78.0), and so is the frustration of being asked to lead what one has not been prepared to do (training, 57.7). Darling-Hammond and Richardson [12], reviewing the evidence on effective teacher professional development, concluded that what works involves workshops accompanied by coaching and mentoring follow-up, with leaders and teachers working regularly in peer learning groups. Sargent and Hannum [58], studying professional learning communities in resource-constrained primary schools in rural China, found that teachers rated highest for teaching skills had initial training supplemented by regular peer lesson planning groups — the combination the Indonesian cascade omits. Hargreaves and Fullan's [25] concept of professional capital points the same way: it is not accumulated through brief socialisation sessions but through social interaction, joint problem-solving, and practice-connected engagement.

The Gresik case mirrors findings from very different contexts. Bradbury et al. [5] found that even English schools given unusual support — funding, released time, expert guidance — remained constrained by the absence of time and relational stability for the interpretive work that curriculum reform requires. Sullanmaa et al. [65] and Salonen-Hakomäki et al. [57] document that Finnish reform success depended on shared sense-making across levels — not on delivery of a finished curriculum text. Chu and Cravens [9], examining principal professional development in China, found that effective programs included locally-based field experts available to coach leaders and provide follow-up support — a form of sustained, practice-connected mentoring absent from the Indonesian socialisation model. In Taiwan, a review of two decades of research found that principals exercised indirect more than direct instructional leadership [49], and principals there faced comparable demands to lead teacher learning ahead of a new curriculum guideline [50]. Al-Daami and Wallace [1] document what happens when curriculum arrives without the interpretive support to enact it: the burden falls at the end of the chain. What distinguishes these cases from Indonesia is not the content of the reform but the professional learning infrastructure built around it — the conditions under which principals and teachers can develop together the working schema that turns a curriculum concept into a lesson.

There is also a methodological lesson. Adding nine survey composites from the same instrument as the outcome raises explained variance to 60 percent. Read as causal explanation, this would be misleading: shared instrumentation, overlapping respondents, and conceptual proximity all contribute, and removing one indicator costs twelve percentage points. The safe reading is narrower but important: whatever explains curriculum management capacity is located in the same domain as leadership practice and professional learning culture, not in the domain of material provision. National assessment data are well suited to showing where a capacity is not located; they are poorly suited, on their own, to showing what builds it. Leithwood et al. [37] made this point at the framework level: measuring structural conditions tells us what a system has; it does not tell us what its educators can do. Systems that read resource indicators as proxies for professional capability will misdirect the support they deploy.

5 Conclusions: a task for local governance

Across 38,579 schools and 488 districts, curriculum management is the weakest of 39 measured principal capacities. More than 85 percent of principals score below adequacy; no district at lower-secondary level, and only one at primary level, reaches an adequate average. The deficit is system-wide, and its distribution — 96 to 98 percent between schools within the same district — points away from geography and district membership and towards a chain of professional learning that ends at each school separately, in the hands of a principal who received a brief socialisation and was then asked to lead what they had barely absorbed.

Principals in Gresik describe that chain failing with precision. They can state what the reform requires. They cannot show a teacher what it looks like in practice. As Fullan [17] argues, durable educational change requires moving beyond technical reform toward the professional conditions that allow change to be genuinely enacted. Twenty-five years of curriculum change — ganti menteri, ganti kurikulum — have produced principals who endorse each reform and cannot operationalise it. The vocabulary has arrived everywhere. The professional learning infrastructure to convert vocabulary into practice has not.

Radinger and Hague [52] closed their review by asking whether, a decade on, Merdeka Belajar would prove to have transformed Indonesian education. The evidence presented here is an early and partial answer for one link in that chain. Four years into the reform, with 76 percent of primary schools already using the new curriculum, the capacity of principals to manage it was the lowest-scoring of 39 measured capacities and fell below adequacy in all but one district at primary level and in every district at lower-secondary level. The reform's design has been judged sound by the most authoritative external review available. What the national assessment adds is that soundness of design and rate of adoption are not evidence of enactment — and that the instrument capable of telling the two apart is already inside the system.

We draw five implications, four for local education governance and one for national assessment.

First, districts should target schools, not themselves. Because nearly all variation lies within districts, ranking districts and intervening in the weakest addresses the wrong unit. School-level National

Assessment results, already returned to every district, allow support to be differentiated by school. What Tolo et al. [72] call intelligent accountability — differentiated, trust-based, school-level support — is structurally closer to what the variance distribution requires.

Second, resource indicators should not be used to identify where support is needed. Facilities, certification and socio-economic composition explain almost none of the variation; a well-provisioned school is as likely as a poor one to have a principal who cannot operationalise the curriculum.

Third, districts already hold the instruments for sustained professional learning but use them for other purposes. Supervisors and principal working groups could carry the coaching, follow-up and peer work that socialisation sessions omit. The professional development literature identifies differentiation, coaching, mentoring follow-up and regular peer learning groups as the features that matter [12, 14, 18], and Sargent and Hannum [58] and Owen and Wong [48] show that such models are achievable in resource-constrained settings. This consensus should not be overstated. Sims and Fletcher-Wood [63] argue that the evidence behind several of its commonly cited features is weaker than the frequency of their citation suggests. Our claim is therefore the narrower one: a single socialisation session lacks every feature that any account of effective professional development proposes, and the Gresik principals identify that absence themselves. That an alternative is feasible within existing structures is shown by one of the districts studied by Arif et al. [3], where a school supervisor asked the principals under his supervision to visit one another's schools and give feedback on what needed improvement. The practice began in 2005 as an initiative of principals and supervisors, not as an instruction from the district head, and became a district programme of peer supervision. It is the kind of school-to-school, practice-connected learning that a socialisation session does not provide, and it required no new institution. This differs in emphasis from the remedy proposed in the existing review literature, which is to establish principal training centres in every province [40]. More training supply may be needed, but a deficit that varies school by school within the same district is unlikely to be closed by provincial provision alone; it calls for support that reaches individual schools. Using these instruments this way requires redirecting some discretionary spending from physical infrastructure to professional learning.

Fourth, the principal's curriculum role should be specified when the curriculum changes. National regulation defines how principals are appointed, not what they do with a curriculum. Where a reform redefines what teachers do with a curriculum, it necessarily redefines what principals must do to support them; leaving that redefinition implicit transfers the interpretive burden to the person least equipped to carry it. Districts that appoint and evaluate principals can make the expectation explicit even where national regulation does not.

Fifth, national assessment systems should collect data on training received and its quality. The Indonesian public release establishes that provision does not explain curriculum management capacity. It cannot establish what does, because the mechanism — the reach, quality, and relational depth of professional learning — is not in the data. Adding items on training receipt, differentiation, follow-up, and

the trust climate of professional development interactions would make the chain itself measurable [8, 44] and give the system, and each district, the instrument it currently lacks.

Our data cannot show that these measures would work. They show where the deficit is not located, and that the level of government formally responsible for schools does not currently differentiate itself on it.

The last mile of curriculum reform is the distance between a principal who can state what a reform requires and one who can show a teacher what it looks like in practice. Provision does not close it. A brief socialisation does not close it. What is most likely to close it is sustained, practice-connected professional learning of the kind that Tikkanen et al. [71], Owen and Wong [48], and the Gresik principals themselves describe, from different vantage points, as what has not yet arrived. The question for local education governance is whether districts can build that infrastructure before the next reform arrives, rather than discovering its absence once the reform is already moving.

AI Declaration

A generative AI tool (Claude, Anthropic) was used to assist with clearity and readibality the manuscript. The authors reviewed and verified all content and take full responsibility for it.

Declarations

Ethics approval

This study was approved by the institutional research ethics committee under the Institute for Research and Community Service (LPPM) of Darul Ulum University, Jombang, Indonesia (Approval No. 135/LPPM-UNDAR/ETIK/XII/2025), in accordance with the Declaration of Helsinki. Freely given, informed consent was obtained from all individual participants involved in this study prior to data collection. The national assessment data were publicly available and de-identified.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Funding

The authors received no specific funding for this work.

Author Contribution

MS and HM conceived the study, analysed the data and wrote the main manuscript text. KN and A.A.Z. collected and coded the qualitative data. All authors reviewed the manuscript.

Data Availability

The National Assessment data (2023 Rapor Publik) are publicly available at https://data.kemendikdasmen.go.id/, with anonymised district codes. The 2025 Neraca Pendidikan Daerah for Gresik, East Java and Indonesia are publicly available at https://npd.kemendikdasmen.go.id/. The survey responses from principals in Gresik are not publicly available because participants were assured anonymity; de-identified data are available from the corresponding author on reasonable request.

References

1. Al-Daami K, Wallace G. Curriculum reform in a global context: a study of teachers in Jordan. J Curric Stud. 2007;39(3):339–60.
2. Amusa H, Mabugu R. The contribution of fiscal decentralisation to regional inequality: empirical results for South African municipalities. ERSA Working Paper 597. Economic Research Southern Africa, Cape Town; 2016.
3. Arif S, Nihayah RW, Usman S, Rarasati N. Does decentralisation promote learning? Local political settlements and education policies in Indonesia. Compare. 2025;55(6):992–1010. https://doi.org/10.1080/03057925.2024.2344523.
4. Bappeda Kabupaten Gresik. Laporan Akhir Penyusunan Indeks Pendidikan Kabupaten Gresik Tahun 2023. Badan Perencanaan Pembangunan, Penelitian dan Pengembangan Daerah Kabupaten Gresik, Gresik; 2024.
5. Bradbury A, Ranken E, Wyse D, Manyukhina Y. Barriers to curriculum reform: the role of context and policy. Br. J. Educ. Stud., advance online publication; 2026. https://doi.org/10.1080/00071005.2026.2652943
6. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol. 2006;3(2):77–101.
7. Braun V, Clarke V. Reflecting on reflexive thematic analysis. Qual Res Sport Exerc Health. 2019;11(4):589–97.
8. Cerna L. Trust: what it is and why it matters for governance and education. OECD Education Working Papers No. 108. OECD Publishing, Paris; 2014. https://doi.org/10.1787/5jxswcg0t6wl-en
9. Chu H, Cravens XC. Principal professional development in China: challenges, opportunities and strategies. Peabody J Educ. 2012;87(2):178–99.

10. Coleman JS, Campbell EQ, Hobson CJ, McPartland J, Mood AM, Weinfeld FD, York RL. Equality of Educational Opportunity. Washington, DC: U.S. Government Printing Office; 1966.
11. Cuban L. How schools change reforms: redefining reform success and failure. Teach Coll Rec. 1998;99(3):453–77.
12. Darling-Hammond L, Richardson N. Research review-teacher learning: what matters? Educ. Leadersh. 2009;66(5):46–53.
13. de Ree J, Muralidharan K, Pradhan M, Rogers H. Double for Nothing? Experimental Evidence on an Unconditional Teacher Salary Increase in Indonesia. Policy Research Working Paper 8264. World Bank, Washington, DC; 2017. https://doi.org/10.1596/1813-9450-8264
14. Desimone LM. Improving impact studies of teachers' professional development: toward better conceptualizations and measures. Educ Res. 2009;38(3):181–99.
15. Ehren M, Paterson A, Baxter J. Accountability and trust: two sides of the same coin? J Educ Change. 2020;21(1):183–213. https://doi.org/10.1007/s10833-019-09352-4.
16. Fan F, Li M, Tao R, Yang D. Transfer-based decentralisation, economic growth and spatial inequality: evidence from China's 2002–2003 tax sharing reform. Urban Stud. 2020;57(4):806–26. https://doi.org/10.1177/0042098019856780.
17. Fullan M. The new meaning of educational change. New York: sixth ed. Teachers College; 2025.
18. Fusarelli BC, Fusarelli LD. Leadership for the future: enhancing principal preparation through standards and innovation. Educ Sci. 2024;14(12):1403.
19. Gonschorek GJ, Schulze GG. Continuity or change? Indonesia's intergovernmental fiscal transfer system under Jokowi. J Southeast Asian Econ. 2018;35(2):143–64.
20. Hallinger P, Murphy J. Assessing the instructional management behavior of principals. Elem Sch J. 1985;86(2):217–47.
21. Hamaniuk VA, Pristupa YV, Semerikov SO. Why educational change accumulates rather than replaces in Ukrainian natural science education from 1950 to 2022. Discov Educ. 2026;5:1128. https://doi.org/10.1007/s44217-026-02247-6.
22. Hanushek EA. Assessing the effects of school resources on student performance: an update. Educ Eval Policy Anal. 1997;19(2):141–64.
23. Hanushek EA. Education production functions. In: Bradley S, Green C, editors. The Economics of Education. second ed. London: Academic; 2020. pp. 161–70.
24. Hargreaves A. Inclusive and exclusive educational change: emotional responses of teachers and implications for leadership. J Educ Change. 2004;5(2):287–309.
25. Hargreaves A, Fullan M. Professional Capital: Transforming Teaching in Every School. New York: Teachers College; 2012.
26. Hayes D. Cascade training and teachers' professional development. ELT J. 2000;54(2):135–45.
27. Hidayat MT, Suryadi S, Latifannisa N, Sari SN. Rino. Evolution of the education curriculum in Indonesia. J Innov Educ Cult Res. 2025;6(2):381–95. https://doi.org/10.46843/jiecr.v6i2.1312.

28. Iqdami MN, Haryono, Astuti T, Widhanarto GP. Indonesian teachers' perceptions of MOOCs for digital professional development. Discov Educ. 2026;5:1080. https://doi.org/10.1007/s44217-026-02170-w.
29. James E, King EM, Suryadi A. Finance, management, and costs of public and private schools in Indonesia. Econ Educ Rev. 1996;15(4):387–98. https://doi.org/10.1016/S0272-7757(96)00035-0.
30. Kemendikbudristek PM, Pendidikan. Kebudayaan, Riset, dan Teknologi Republik Indonesia Nomor 40 Tahun 2021 tentang Penugasan Guru sebagai Kepala Sekolah [Regulation of the Minister of Education, Culture, Research, and Technology No. 40 of 2021 on the Assignment of Teachers as School Principals]. Kebudayaan, Riset, dan Teknologi, Jakarta: Kementerian Pendidikan; 2021.
31. Kemendikdasmen. Neraca Pendidikan Daerah 2025: Kabupaten Gresik. Kementerian Pendidikan Dasar dan Menengah, Jakarta; 2025.
32. Kemendikdasmen. Neraca Pendidikan Daerah 2025: Provinsi Jawa Timur. Kementerian Pendidikan Dasar dan Menengah, Jakarta; 2025.
33. Kemendikdasmen. Neraca Pendidikan Daerah 2025: Indonesia. Jakarta: Kementerian Pendidikan Dasar dan Menengah; 2025.
34. Kemendikdasmen. Naskah Akademik Pembelajaran Mendalam. Jakarta: Kementerian Pendidikan Dasar dan Menengah; 2025.
35. Kemendikdasmen. Peraturan Menteri Pendidikan Dasar dan Menengah Republik Indonesia Nomor 7. Tahun 2025 tentang Penugasan Guru sebagai Kepala Sekolah [Regulation of the Minister of Primary and Secondary Education No. 7 of 2025 on the Assignment of Teachers as School Principals]. Kementerian Pendidikan Dasar dan Menengah, Jakarta; 2025.
36. Khairina N, Yarrow N, Cilliers J, Dini I. Improving Teachers and School Leadership in Indonesia: Impact Evaluation of Guru Penggerak Program at the Primary Level. Jakarta: World Bank; 2023.
37. Leithwood K, Jantzi D, Mascall B. A framework for research on large-scale reform. J Educ Change. 2002;3(1):7–33.
38. Lewis BD, Smoke P. Intergovernmental fiscal transfers and local incentives and responses: the case of Indonesia. Fisc Stud. 2017;38(1):111–39. https://doi.org/10.1111/1475-5890.12080.
39. Louis KS. Trust and improvement in schools. J Educ Change. 2007;8(1):1–24. https://doi.org/10.1007/s10833-006-9015-5.
40. Lumban Gaol NT. School leadership in Indonesia: a systematic literature review. Educ. Manag. Adm. Leadersh., advance online publication; 2021. https://doi.org/10.1177/17411432211010811
41. Moos L. How do schools bridge the gap between external demands for accountability and the need for internal trust? J Educ Change. 2005;6(4):307–28. https://doi.org/10.1007/s10833-005-2749-7.
42. Muttaqin T, van Duijn M, Heyse L, Wittek R. The impact of decentralisation on educational attainment in Indonesia. In: Holzhacker RL, Wittek R, Woltjer J, editors. Decentralization and Governance in Indonesia. Cham: Springer; 2016. pp. 79–103. https://doi.org/10.1007/978-3-319-22434-3_4.

43. Neumerski CM. Rethinking instructional leadership, a review: what do we know about principal, teacher, and coach instructional leadership, and where should we go from here? Educ Adm Q. 2013;49(2):310–47.
44. Niedlich S, Kallfaß A, Pohle S, Bormann I. A comprehensive view of trust in education: conclusions from a systematic literature review. Rev Educ. 2021;9:124–58. https://doi.org/10.1002/rev3.3239.
45. Nurmuhaemin W. Perubahan kurikulum Indonesia butuh 'committee of the future'. Kompas: 15 January; 2023.
46. OECD. Education Policy Outlook 2015: Making Reforms Happen. Paris: OECD Publishing; 2015.
47. OECD, Curriculum Reform. A Literature Review to Support Effective Implementation. OECD Education Working Papers No. 239. OECD Publishing, Paris; 2020.
48. Owen S, Wong D. Timor-Leste education: supporting sustainable system-wide reform and school leader capacity-building through collaborative foreign aid. J Educ Change. 2021;22(3):379–400. https://doi.org/10.1007/s10833-020-09397-w.
49. Pan HLW, Nyeu FY, Chen JS. Principal instructional leadership in Taiwan: lessons from two decades of research. J Educ Adm. 2015;53(4):492–511. https://doi.org/10.1108/JEA-01-2014-0006.
50. Pan HLW, Nyeu FY, Cheng SH. Leading school for learning: principal practices in Taiwan. J Educ Adm. 2017;55(2):168–85. https://doi.org/10.1108/JEA-06-2016-0069.
51. Priestley M, Alvunger D, Philippou S, Soini T, editors. Curriculum Making in Europe. Bingley: Emerald; 2021.
52. Radinger T, Hague C. Transforming Education in Indonesia: Examining the Landscape of Current Reforms. OECD Education Policy Perspectives No. 88. OECD Publishing, Paris; 2023. https://doi.org/10.1787/9ff8d407-en
53. Republik Indonesia. Undang-Undang Nomor 23 Tahun 2014 tentang Pemerintahan Daerah [Law No. 23 of 2014 on Regional Government]. Jakarta; 2014.
54. Revina S, Pramana RP, Fillaili R, Suryadarma D. Systemic Constraints Facing Teacher Professional Development in a Middle-Income Country. RISE Working Paper 20/054. Oxford; 2020.
55. Sahlberg P. Trends in global education reform movement since the 1990s. Int J Educ Dev. 2023;98:3–13. https://doi.org/10.1016/j.ijedudev.2023.102748.
56. Sahlberg P. Trust: the missing link in educational change. J Educ Change. 2026. https://doi.org/10.1007/s10833-026-09568-1.
57. Salonen-Hakomäki SM, Soini T, Pietarinen J, Pyhältö K. Leading complex educational change via national participative reforms? J Educ Change. 2024;25(3):531–54.
58. Sargent TC, Hannum E. Doing more with less: teacher professional learning communities in resource-constrained primary schools in rural China. J Teach Educ. 2009;60(3):258–76.
59. Saunders R. The role of teacher emotions in change: experiences, patterns and implications for professional development. J Educ Change. 2013;14(3):303–33.

60. Shaked H. Instructional leadership clarity: how principals make instructional leadership intelligible and predictable for teachers. J. Educ. Change, advance online publication; 2026.
61. Shaoan MMR. Principals' instructional leadership on teachers' instructional practices in rural primary schools in Sirajganj. Bangladesh Sage Open. 2026;16(1). https://doi.org/10.1177/21582440261417006.
62. Sholeh M, Windasari, Lestari DN, Widodo BS. Selection process of school principals in Indonesia and some other countries. J Penelit Pengemb Pendidik. 2023;7(1):88–95. https://doi.org/10.23887/jppp.v7i1.56174.
63. Sims S, Fletcher-Wood H. Identifying the characteristics of effective teacher professional development: a critical review. Sch Eff Sch Improv. 2021;32(1):47–63.
64. Spillane JP, Reiser BJ, Reimer T. Policy implementation and cognition: reframing and refocusing implementation research. Rev Educ Res. 2002;72(3):387–431.
65. Sullanmaa J, Pyhältö K, Pietarinen J, Soini T. Curriculum coherence as perceived by district-level stakeholders in large-scale national curriculum reform in Finland. Curric J. 2019;30(3):244–63.
66. Sumintono B, Sheyoputri EYA, Jiang N, Misbach IF, Jumintono. Becoming a principal in Indonesia: possibility, pitfalls and potential. Asia Pac J Educ. 2015;35(3):342–52.
67. Sumintono B, Hariri H, Nugroho PJ, Izzati UA, Sriyanto J. The role of districts in supporting school leaders' instructional leadership: a view and experience from a developing country. J Educ Adm. 2019;57(5):591–600. https://doi.org/10.1108/JEA-09-2019-227.
68. Sumiyana S, Effendi R. Indonesia's education budgeting system denoting low adaptiveness: functionalism analysis. Cogent Soc Sci. 2024;10(1):2393870. https://doi.org/10.1080/23311886.2024.2393870.
69. Talitha T, Firman T, Hudalah D. Welcoming two decades of decentralisation in Indonesia: a regional development perspective. Territ Polit Gov. 2020;8(5):690–708. https://doi.org/10.1080/21622671.2019.1601595.
70. Thoonen EEJ, Sleegers PJC, Oort FJ, Peetsma TTD, Geijsel FP. How to improve teaching practices: the role of teacher motivation, organizational factors, and leadership practices. Educ Adm Q. 2011;47(3):496–536.
71. Tikkanen L, Pyhältö K, Pietarinen J, Soini T. Lessons learnt from a large-scale curriculum reform: the strategies to enhance development work and reduce reform-related stress. J Educ Change. 2019;20(4):487–513. https://doi.org/10.1007/s10833-019-09363-1.
72. Tolo A, Lillejord S, Flórez Petour MT, Hopfenbeck TN. Intelligent accountability in schools. J Educ Change. 2020;21(1):59–82. https://doi.org/10.1007/s10833-019-09359-x.
73. Vernez G, Karam R, Marshall JH. Implementation of School-Based Management in Indonesia. Santa Monica, CA: RAND Corporation; 2012.
74. Weiner BJ. A theory of organizational readiness for change. Implement Sci. 2009;4:67.

Figures

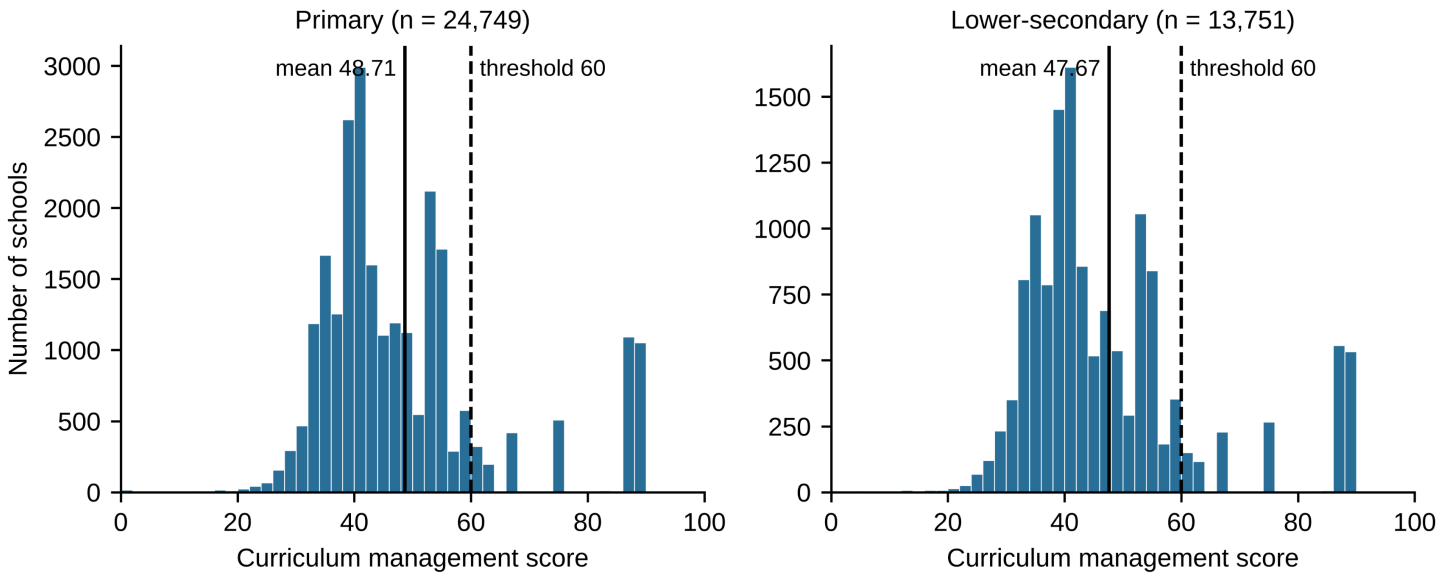


Figure 1

Distribution of curriculum management scores, National Assessment 2023. Solid line = sample mean; dashed line = adequacy threshold of 60.

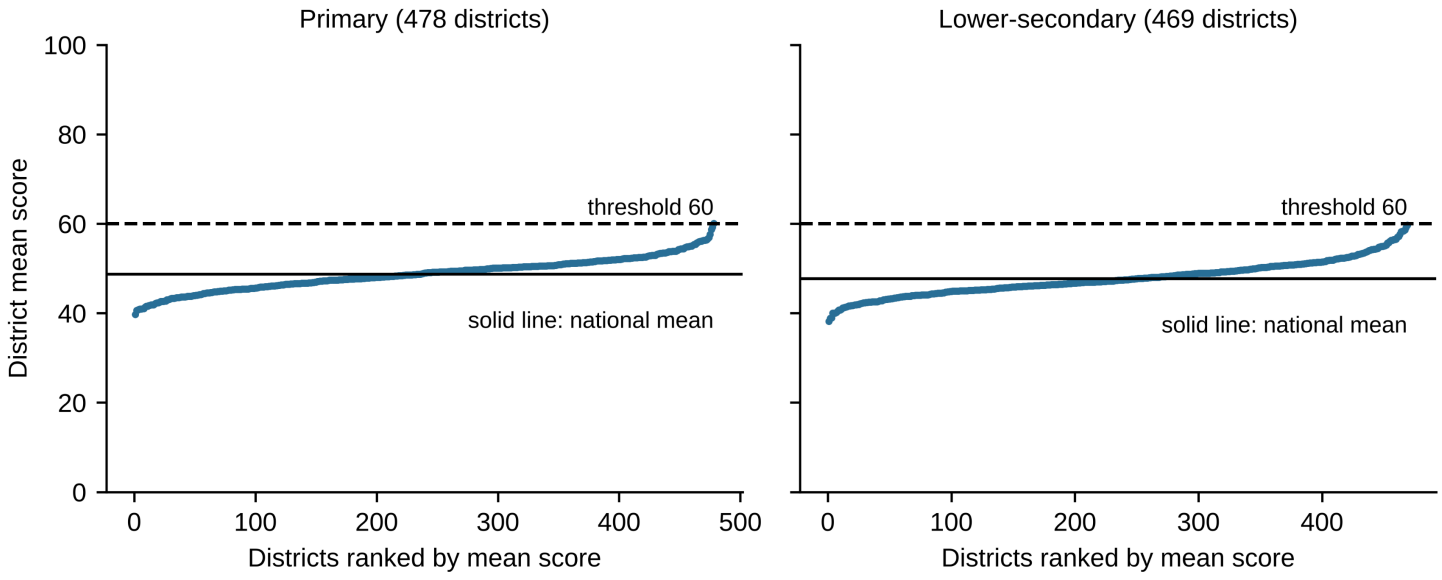


Figure 2

District and city means, ranked from lowest to highest. Districts with fewer than ten sampled schools excluded.

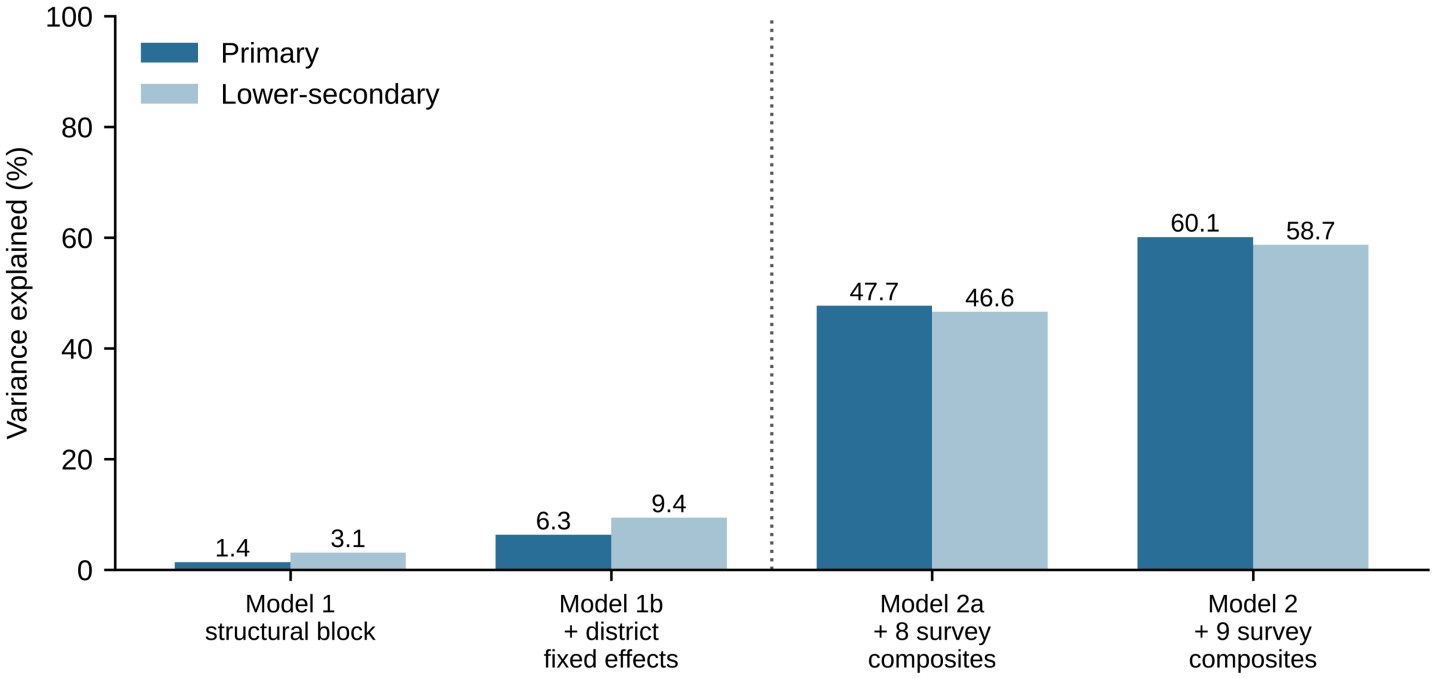


Figure 3

Variance in curriculum management explained, by model. Blocks right of the dotted line include same-instrument composites; their evidentiary status differs from the structural models on the left.

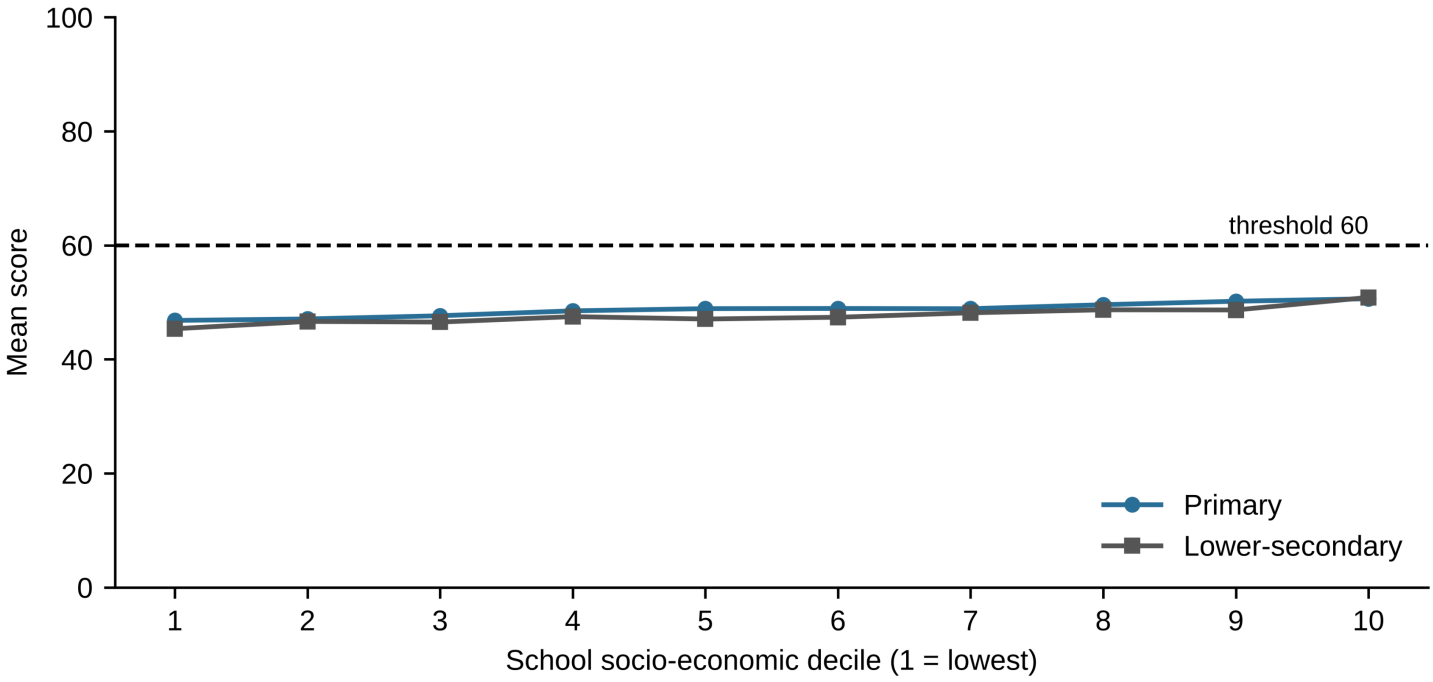


Figure 4

Mean curriculum management score by school socio-economic decile. Vertical axis covers the full scale to show both the flatness of the gradient and the distance from the threshold.