Berkala Ilmiah Pendidikan
Volume 6, Number 3, November 2026

scidac plus

Testing The Improvement of Deep Learning Quality Viewed from Principals' Visionary Leadership and Teachers' Professional Competence In Elementary Schools

Lena Rhosita Zaenal[1], Sri Utaminingsih[2], Edris Zamroni[3*]
[1,2,3]Universitas Muria Kudus, Kudus, Indonesia
edris.zamroni@umk.ac.id*

Abstract

This study examined how principals' visionary leadership and teachers' professional competence are associated with the quality of deep learning in elementary schools. A cross-sectional quantitative survey was conducted among 142 teachers from 31 public elementary schools in Donorojo District, Jepara, Indonesia. Hypotheses were tested with partial least squares structural equation modelling (PLS-SEM) using 5,000 bootstrap resamples, after evaluating indicator reliability, internal consistency, convergent and discriminant validity, collinearity, and common method bias. Visionary leadership (β = 0.256, p = .001, f^2 = 0.12) and professional competence (β = 0.649, p < .001, f^2 = 0.78) both had positive and significant effects on deep learning quality, jointly explaining 71.6% of its variance (adjusted R^2 = 0.711; SRMR = 0.076). A supplementary analysis indicated that part of the leadership effect operated through teachers' professional competence (indirect effect = 0.447, 95% CI [0.337, 0.561]). However, discriminant validity between professional competence and deep learning quality was not fully established (HTMT = 0.955), and school-level observation scores did not correlate with teacher-reported scores. The findings suggest that a principal's vision matters most when translated into teachers' professional capability, but they should be interpreted cautiously because all constructs were self-reported through a single instrument.

Keywords: visionary leadership, teacher professional competence, deep learning, PLS-SEM, elementary school

INTRODUCTION

Indonesia's Ministry of Primary and Secondary Education introduced the deep learning approach in 2025 through an academic manuscript that frames quality learning around three principles: mindful, meaningful, and joyful learning (Kemendikdasmen, 2025). The approach is presented as a way of changing how teaching is done rather than replacing the curriculum, and it is consistent with constructivist views in which learners actively construct understanding (Rosyidah et al., 2026). The idea has older roots in the distinction between deep and surface approaches to learning (Marton & Säljö, 1976), in evidence that the effectiveness of learning strategies differs between surface and deep phases of learning (Hattie & Donoghue, 2016), and in work on new pedagogies that build learning partnerships between teachers and students (Fullan & Langworthy, 2014). Studies of schools also suggest that deeper learning is valued more often than it is realized in ordinary classrooms (Mehta & Fine, 2019). For primary schools, the practical question is therefore which school-level and teacher-level conditions make deep learning more likely.

Artikel ini menggunakan lisensi Creative Commons Attribution 4.0 International License
© Lena Rhosita Zaenal et al. (2026)

Two conditions are repeatedly highlighted in the literature. The first is school leadership. A large body of research indicates that leadership affects learning mainly indirectly, by shaping the organizational conditions, staff motivation, and working conditions that support teaching (Hallinger & Heck, 1998; Leithwood et al., 2008, 2020). Reviews of leadership research also note that leadership models differ in emphasis (Bush & Glover, 2014) and that leaders' promotion of and participation in teacher learning is among the dimensions most strongly associated with student outcomes (Robinson et al., 2008). Reviews spanning several decades of research further support the view that leadership matters for learning through carefully specified mechanisms rather than as a generic trait (Hallinger, 2011). Visionary leadership, which concerns formulating, communicating, and enacting a shared direction for the school, is one of the leadership orientations most discussed in Indonesian school research. For example, a study of elementary school teachers in Palembang reported a positive and significant effect of visionary leadership on teacher professionalism (Lian et al., 2025), and a path analysis in early childhood education reported a positive effect of visionary leadership on teacher performance (Kusumawati, 2023).

The second condition is teachers' professional competence. Shulman's (1986, 1987) work on the knowledge base of teaching established pedagogical content knowledge as central to professional competence, and later work extended this to the integration of technology with content and pedagogy (Mishra & Koehler, 2006). Empirically, teachers' professional competence has been linked to instructional quality and, through it, to student outcomes (Baumert et al., 2010; Kunter et al., 2013). Comparative research on teacher education argues that high-performing systems invest systematically in teachers' professional learning (Darling-Hammond, 2017), and Indonesian scholarship has discussed the demands that twenty-first-century learners place on teachers' pedagogical competence (Ofita & Sururi, 2023). For deep learning specifically, Indonesian studies have emphasized the need to build teachers' deep-learning mindset and to integrate the approach into everyday practice (Hendrianty et al., 2024; Rahmandani et al., 2025), and systematic reviews are beginning to map how it is understood (Akmal et al., 2025).

Despite these strands of evidence, studies that estimate the effects of principals' visionary leadership and teachers' professional competence on deep learning quality within one structural model remain scarce in the sources we reviewed, particularly for elementary schools in Indonesia where the approach is new. Estimating both predictors together matters because leadership and competence are likely to be correlated, so single-predictor studies may over- or under-state each contribution. This study therefore estimates a structural equation model to (1) test the effect of visionary leadership (X1) on deep learning quality (Y), (2) test the effect of teachers' professional competence (X2) on deep learning quality, and (3) assess how much of the variance in deep learning quality both predictors explain together. The following hypotheses were tested: H1: visionary leadership has a positive effect on deep learning quality; H2: teachers' professional competence has a positive effect on deep learning quality; H3: visionary leadership and professional competence jointly explain a substantial proportion of the variance in deep learning quality. An exploratory mediation path (X1 $\longrightarrow$ X2 $\longrightarrow$ Y) was examined as a supplementary analysis because the two predictors were expected to be correlated.

METHOD

Research design, setting, and participants

This study used a cross-sectional, quantitative survey design with hypothesis testing by partial least squares structural equation modelling (PLS-SEM). It was conducted in public elementary schools (SD Negeri) in Donorojo District, Jepara Regency, Central Java, Indonesia. Data were collected online between 12 and 19 August 2026. A total of 142 teachers from 31 schools (three to six respondents per school) completed the questionnaire. This study employs a proportionate random sampling technique based on probability sampling

to ensure equal opportunity and proportional representation for all public elementary school teachers in Donorojo District, Jepara Regency (Sugiyono, 2022). The minimum sample size was determined using Slovin's formula with a 5% margin of error (Surani et al., 2023). Combining Slovin's formula and proportionate random sampling ensures both quantitative adequacy and sample distribution quality, making the collected data representative and allowing the findings to be generalized to the target population (Daniar Pramita et al., 2021). Supplementary information was obtained from school supervisors of three supervisory regions (Dabin) and from school observation sheets covering the same 31 schools; these data were used only for descriptive triangulation and were not part of the structural model.

Instruments and Variable Measurement

The teacher questionnaire contained 20 statements rated on a five-point scale (1 to 5). Items 1–8 measured principals' visionary leadership (X1), items 9–14 measured teachers' professional competence (X2), and items 15–20 measured deep learning quality (Y). All three constructs were modelled reflectively. Visionary leadership refers to the principal's capacity to formulate, communicate, and enact a shared vision for school improvement; professional competence refers to teachers' mastery of content, pedagogy, and professional practice; and deep learning quality refers to the extent to which classroom learning is mindful, meaningful, and joyful (Kemendikdasmen, 2025).

Instrument Specification Grid

Variable	Dimension/Indicator	Sub-Indicator	Item No.	Type (+/-)	Item Count	Source/Adaptation
Principal Visionary Leadership (X_1)	Thinking about the school's future	Planning school development to achieve the school's vision	1	+	1	Adapted from Visionary Leadership Theory
	Creating an advanced & anticipatory organizational culture	Building a school culture that is adaptive to change and innovation	2	-	1	
	Realizing a quality school	Developing continuous school quality improvement programs	3	+	1	
	Clarifying high aspirations	Communicating the school's vision and goals clearly to all school members	4	-	1	
	Fostering inspiration & commitment	Encouraging the commitment of school members in realizing the school's vision	5	+	1	
	Conveying organizational values	Instilling school values in every program and activity execution	6	-	1	

Variable	Dimension/Indicator	Sub-Indicator	Item No.	Type (+/-)	Item Count	Source/Adaptation
	Motivating in the right direction	Providing motivation to school members to achieve common goals	7	+	1	
	Coordinating diverse abilities	Optimizing the potential of each school member according to their abilities and roles	8	-	1	
Teacher Professional Competence (X_2)	Mastering subject matter & scientific mindset	Mastering concepts and learning materials in accordance with the subject taught	9	+	1	Adapted from Teacher Professional Competence Standards
	Mastering learning outcomes & indicators	Understanding learning outcomes and learning objectives as the foundation for learning execution	10	-	1	
	Developing material creatively	Developing learning materials according to student characteristics and needs	11	+	1	
	Acting reflectively & continuously	Conducting reflections to continuously improve learning quality	12, 13	-, +	2	
	Utilizing Information Technology (ICT)	Using information technology to support the learning process	14	-	1	
Quality of Deep Learning (Y)	Active student engagement in learning	Active participation of students in every learning stage	15	+	1	Adapted from Deep Learning Framework & 21st Century Skills (4Cs/HOTS)
	Implementation of student-centered learning strategies	Utilizing learning strategies that provide opportunities for students to construct knowledge independently	16	-	1	
	Development of critical, creative, collaborative, and communicative thinking skills	Providing learning experiences that develop Higher-Order Thinking Skills (HOTS) and 21st-century skills (4Cs)	17	+	1	
	Relevance of learning materials to real-world contexts	Integrating learning materials with real-life problems or experiences relevant to students' lives	18	-	1	

Variable	Dimension/Indicator	Sub-Indicator	Item No.	Type (+/-)	Item Count	Source/Adaptation
	Achievement of learning objectives and competencies	Attainment of learning objectives in accordance with established learning outcomes	19, 20	+, -	2	
Total Items					20	

Content validity was evaluated through expert judgment by specialists in Educational Management and Research Methodology. The evaluation focused on the alignment between item statements and indicators, language clarity, and construct representation. Feedback and recommendations from the experts served as the basis for revising and refining the instrument prior to empirical field testing. The supervisor questionnaire had 16 items (maximum score 80 per supervisory region), and the observation sheet had 15 items (score range 15–75).

Data screening and analysis

Four of 2,840 item responses (0.14%) were missing, all on deep-learning items from three respondents; each was replaced with the respondent's mean on the remaining items of the same scale. One pair of identical records from the same school was found; removing one record (n = 141) left the results practically unchanged (β = 0.261 and 0.648; R^2 = 0.721), so all 142 records were retained. Univariate and multivariate normality were checked with Shapiro–Wilk and Mardia's tests. Because the data were non-normal and the model has a prediction-oriented purpose, PLS-SEM was chosen as the estimation method (Hair et al., 2019, 2022). Sample adequacy was evaluated with a power analysis for a regression with two predictors (Cohen, 1988): a minimum of 68 cases is required to detect a medium effect (f^2 = 0.15) with power of .80 at α = .05, and the achieved power with n = 142 was 0.99.

Estimation used the PLS algorithm with the path-weighting scheme and Mode A outer weighting, implemented in Python (NumPy/SciPy) following the standard procedure described by Hair et al. (2019, 2022) and Henseler et al. (2016). Significance was assessed with 5,000 bootstrap resamples (percentile 95% confidence intervals, two-tailed). The measurement model was evaluated through indicator loadings ($\geq$ .708 preferred; .40–.70 retained when AVE and composite reliability criteria are met), Cronbach's alpha, Dijkstra–Henseler's ρA (Dijkstra & Henseler, 2015) and composite reliability (.70–.95), average variance extracted (AVE $\geq$ .50), the Fornell–Larcker criterion (Fornell & Larcker, 1981), cross-loadings, and the heterotrait–monotrait ratio (HTMT; Henseler et al., 2015; threshold .85/.90), with outer variance inflation factors (VIF < 3). Common method bias was examined with Harman's single-factor test and the full-collinearity VIF (threshold 3.3; Kock, 2015), following Podsakoff et al. (2012). The structural model was evaluated through inner VIF, path coefficients, R^2, effect size f^2 (.02, .15, and .35 as small, medium, and large; Cohen, 1988), standardized root mean square residual (SRMR < .08), and out-of-sample predictive power using ten-fold cross-validation repeated five times (Q^2predict and comparison with a linear-model benchmark; Shmueli et al., 2019). As robustness checks, a school-level (cluster) bootstrap with 2,000 resamples was run because teachers were nested in schools, and intraclass correlations were computed.

RESULTS AND DISCUSSION

Descriptive statistics and data screening

Table 1 summarizes the three constructs. Teachers rated all constructs highly: the percentage of the ideal score was 90.72% for visionary leadership, 88.57% for professional competence, and 89.67% for deep learning quality, and no respondent fell into the moderate, low, or very low categories. Item means ranged from 4.33 to 4.63 on the five-point scale and responses were concentrated on the values 3 to 5. Supervisors in the three Dabin rated visionary leadership between 90.00% and 92.50% (mean 91.25%) on their 16-item instrument, which is in line with the teachers' ratings. School observation scores averaged 59.90 (SD = 4.98; range 53–71), with 10 schools classified as very consistent and 21 as consistent.

Table 1. Descriptive statistics of the study constructs (n = 142)

Construct	Items	Item mean	Scale sum, M (SD)	% of ideal score	Very high / High (n)
Visionary leadership (X1)	8	4.54	36.29 (3.00)	90.72	113 / 29
Professional competence (X2)	6	4.43	26.57 (2.30)	88.57	99 / 43
Deep learning quality (Y)	6	4.48	26.90 (2.78)	89.67	97 / 45

Note. Categories follow the 81–100% (very high) and 61–80% (high) intervals; no respondent fell below 61%. Deep-learning scale sums use mean-imputed values for four missing responses.

Normality checks indicated clear departures from normality. All Shapiro–Wilk tests were significant ($p < .001$), item skewness ranged from -0.64 to 0.32, and Mardia's multivariate skewness (101.95, $\chi^2 = 2412.7$, df = 1540, $p < .001$) and kurtosis (497.12, $z = 11.47$, $p < .001$) were both significant. This ceiling-type distribution supports the use of PLS-SEM, which does not assume multivariate normality, but it also implies limited variability that can attenuate correlations.

Measurement model assessment

Table 2 reports the indicator-level results. Outer loadings ranged from 0.550 to 0.846, and all were statistically significant (minimum $t = 8.33$, $p < .001$). Fifteen of the 20 indicators met the preferred threshold of .708. Five indicators (items 1, 4, 5, 9, and 14) had loadings between .55 and .71; they were retained because the AVE and composite reliability of their constructs exceeded the recommended cut-offs (Hair et al., 2019). Outer VIF values were all below 3 (maximum = 2.51), indicating no problematic collinearity among indicators.

Table 2. Indicator loadings, bootstrap significance, and outer VIF

Indicator	Construct	Mean	Loading	t	p	Outer VIF
VL 1	X1	4.63	0.641	8.76	< .001	1.90
VL 2	X1	4.61	0.720	11.95	< .001	2.23
VL 3	X1	4.51	0.780	22.36	< .001	1.93
VL 4	X1	4.48	0.671	12.56	< .001	1.53
VL 5	X1	4.44	0.631	12.20	< .001	1.47
VL 6	X1	4.54	0.726	15.01	< .001	1.74
VL 7	X1	4.58	0.804	19.56	< .001	2.42
VL 8	X1	4.50	0.791	22.95	< .001	2.21
PC 9	X2	4.33	0.550	8.33	< .001	1.27
PC 10	X2	4.55	0.798	21.49	< .001	2.05
PC 11	X2	4.45	0.799	19.67	< .001	2.00
PC 12	X2	4.42	0.764	19.71	< .001	1.83
PC 13	X2	4.48	0.797	19.84	< .001	2.01
PC 14	X2	4.35	0.630	12.10	< .001	1.41
DL1	Y	4.50	0.846	30.34	< .001	2.51
DL2	Y	4.49	0.818	25.49	< .001	2.18

Indicator	Construct	Mean	Loading	t	p	Outer VIF
DL3	Y	4.56	0.817	22.62	< .001	2.29
DL4	Y	4.44	0.828	27.01	< .001	2.25
DL5	Y	4.49	0.812	23.25	< .001	2.22
DL6	Y	4.42	0.833	26.04	< .001	2.48

Note. t and p values are based on 5,000 bootstrap resamples.

Table 3 shows that all constructs were reliable and convergently valid. Cronbach's alpha ranged from 0.819 to 0.906, ρA from 0.832 to 0.907, and composite reliability from 0.870 to 0.928, all within the .70–.95 range. AVE values (0.523, 0.532, and 0.682) exceeded .50. Thus, reliability and convergent validity were established.

Table 3. Construct reliability and convergent validity

Construct	Items	Cronbach's α	ρA	CR	AVE
Visionary leadership (X1)	8	0.869	0.879	0.897	0.523
Professional competence (X2)	6	0.819	0.832	0.870	0.532
Deep learning quality (Y)	6	0.906	0.907	0.928	0.682

Discriminant validity results are presented in Table 4. Under the Fornell–Larcker criterion, the square root of AVE exceeded the inter-construct correlations for X1 and Y, but not for X2 ($\sqrt{\text{AVE}}$ = 0.730 versus r(X2, Y) = 0.825). The square root of AVE for Y (0.826) was virtually identical to its correlation with X2 (0.825). The HTMT ratios for X1–X2 (0.795) and X1–Y (0.771) were below .85, but the ratio for X2–Y (0.955) exceeded the conservative .90 threshold and its bootstrap confidence interval reached 1 (upper 95% bound = 1.018). Cross-loadings, in contrast, were satisfactory, with every indicator loading highest on its own construct. In short, X1 is empirically distinct from the other constructs, whereas professional competence and deep learning quality overlap strongly; discriminant validity between these two constructs is therefore not fully established, and the X2 ⟶ Y path should be read with this caveat.

Table 4. Discriminant validity: Fornell–Larcker criterion and HTMT

Pair	Latent r	$\sqrt{\text{AVE}}$ (first / second)	$\sqrt{\text{AVE}}$ > r for both?	HTMT [95% CI]	Decision
X1 – X2	0.690	0.723 / 0.730	Yes	0.795 [0.668, 0.906]	Established
X1 – Y	0.704	0.723 / 0.826	Yes	0.771 [0.656, 0.877]	Established
X2 – Y	0.825	0.730 / 0.826	No	0.955 [0.891, 1.018]	Not established

Note. HTMT confidence intervals are percentile intervals from 2,000 bootstrap resamples. Decision is based on both criteria and the HTMT thresholds of .85 and .90.

Common method bias was also examined because all constructs came from one self-report instrument. Harman's single-factor test showed that the first unrotated factor explained 47.1% of the variance, below the 50% benchmark. However, full-collinearity VIFs were 2.14 (X1), 3.39 (X2), and 3.52 (Y); the values for X2 and Y exceed the 3.3 threshold suggested by Kock (2015), which signals possible common method bias or strong construct overlap. Method variance therefore cannot be ruled out and is acknowledged as a limitation (Podsakoff et al., 2012).

Structural model assessment and hypothesis testing

Collinearity between the predictors was acceptable (inner VIF = 1.91, with X1 and X2 correlated at r = 0.690). Table 5 reports the structural results. Visionary leadership had a positive and significant effect on

deep learning quality (β = 0.256, t = 3.34, p = .001, 95% CI [0.117, 0.415]), supporting H1. Professional competence also had a positive and significant effect (β = 0.649, t = 9.11, p < .001, 95% CI [0.502, 0.779]), supporting H2. Competence was the stronger predictor: its effect size was large (f^2 = 0.78), whereas the effect size of leadership was small to medium (f^2 = 0.12, just below the .15 benchmark). Together, the two predictors explained 71.6% of the variance in deep learning quality (R^2 = 0.716, adjusted R^2 = 0.711, 95% CI [0.655, 0.798]), a substantial level, which supports H3 (Figure 1).

Table 5. Structural model results and hypothesis testing

Hypothesis / path	β (R^2)	SE	t	p	95% CI	f^2	Decision
H1: X1 ⟶ Y	0.256	0.077	3.34	.001	[0.117, 0.415]	0.121	Supported
H2: X2 ⟶ Y	0.649	0.071	9.11	< .001	[0.502, 0.779]	0.775	Supported
H3: R^2 of Y	0.716				[0.655, 0.798]		Supported

Note. Bootstrap with 5,000 resamples; two-tailed tests. Adjusted R^2 = 0.711.

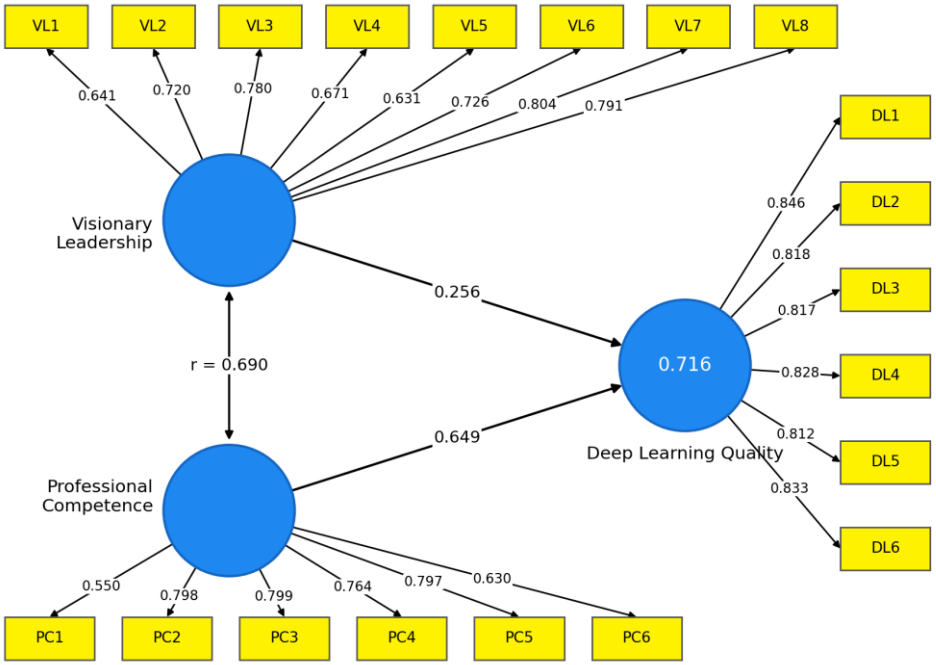


Figure 1. PLS-SEM results: outer loadings (on indicator arrows), standardized path coefficients (structural arrows), and R^2 (inside the endogenous construct)

Model quality indicators supported the structural results. SRMR was 0.076, below the .08 criterion. In the out-of-sample (ten-fold, five repetitions) assessment, Q^2predict for the six deep-learning indicators ranged from 0.418 to 0.550, all above zero; the PLS model produced lower prediction errors than the linear-model benchmark for five of the six indicators and an equal error for the sixth, indicating medium-to-high predictive power (Shmueli et al., 2019). Because teachers were nested within schools (intraclass correlations of 0.22 for X1, 0.18 for X2, and 0.24 for Y), a school-level bootstrap was used as a robustness check. It gave larger standard errors (SE = 0.090 for X1 ⟶ Y and 0.084 for X2 ⟶ Y) but the same conclusions: the 95% confidence intervals were [0.108, 0.455] and [0.462, 0.786], both excluding zero.

Supplementary analyses

Because visionary leadership and professional competence were correlated (r = 0.690, 95% CI [0.584, 0.784]), an exploratory mediation path was estimated from the composite scores of the main model. The indirect effect of visionary leadership on deep learning quality through professional competence was 0.447 (SE = 0.058, 95% CI [0.337, 0.561]); the direct effect was 0.256, and the total effect was 0.704 (95% CI [0.612, 0.795]). About 64% of the total effect therefore operated through competence, indicating partial (complementary) mediation. Given the cross-sectional design, this result describes a pattern consistent with mediation rather than demonstrating a causal sequence.

As a descriptive triangulation, school-level mean teacher scores were compared with the school observation scores (n = 31 schools). The correlations were close to zero for visionary leadership (Spearman ρ = -0.008, p = 0.968), professional competence (ρ = 0.065, p = 0.727), and deep learning quality (ρ = 0.001, p = 0.998). Thus, teacher-reported scores did not track the observed consistency of practice across schools. This may reflect the small and uneven number of respondents per school, the compressed range of both measures, and differences between perceptions and observed behaviour; it also means that the high self-reported levels should not be read as proof of high observed quality.

Discussion

The first main finding is that professional competence was the dominant correlate of deep learning quality (β = 0.649, f^2 = 0.78). This pattern accords with research in which teachers' professional competence predicts instructional quality and, through it, student outcomes (Baumert et al., 2010; Kunter et al., 2013). Teaching for mindful, meaningful, and joyful learning requires a rich knowledge base about content and how to teach it (Shulman, 1986, 1987), and the ability to combine that knowledge with technology where it helps (Mishra & Koehler, 2006). It also fits the argument that systems that develop teachers' capacity are better placed to realize new learning goals (Darling-Hammond, 2017) and the emphasis of Indonesian deep-learning studies on building teachers' mindset and practice (Hendrianty et al., 2024; Rahmandani et al., 2025). Two cautions apply. First, the high HTMT between competence and deep learning quality shows that teachers may not have distinguished clearly between what they know and how well learning unfolds in their classes, so the size of this path is probably inflated by construct overlap and shared method variance. Second, the competence items were self-assessed, and self-assessment tends to be generous.

The second finding is that visionary leadership had a smaller but significant direct effect (β = 0.256, f^2 = 0.12) on deep learning quality. The smaller direct coefficient is consistent with the view that leadership works largely indirectly, by shaping staff motivation, commitment, and working conditions (Hallinger & Heck, 1998; Leithwood et al., 2008, 2020), and that leaders have the strongest influence when they promote and take part in teacher learning (Robinson et al., 2008). The supplementary analysis points in the same direction: roughly 64% of the total effect of leadership ran through professional competence. In practical terms, a vision for deeper learning appears to matter most where it is turned into opportunities for teachers to build their competence, for example through collaborative learning among teachers, supervision, and coaching. Indonesian studies reporting positive effects of visionary leadership on teacher professionalism and performance (Kusumawati, 2023; Lian et al., 2025) are compatible with this reading, although those studies used different outcomes and, in one case, a much smaller effect.

Third, the conceptual framing of the outcome helps explain why the findings are plausible. Deep learning is not simply more learning; it concerns the qualitative way in which learners engage with meaning (Marton & Säljö, 1976) and the match between strategies and the phase of learning (Hattie & Donoghue,

2016). It is associated with open and context-sensitive attention (Langer, 2016; Bordunos et al., 2024), with the connection of new knowledge to existing understanding (Mystakidis, 2021), and with joyful engagement as emphasized in the national framework (Kemendikdasmen, 2025; Rosyidah et al., 2026). Because these qualities are produced in classroom interaction, they depend on teachers' daily decisions more than on distant leadership statements, which again supports the stronger role of competence. At the same time, the observation that deeper learning is often rarer in practice than in aspiration (Mehta & Fine, 2019) is a reminder to treat uniformly high self-reports with care.

The study has implications for school improvement. School leaders in this district can strengthen deep learning most effectively by converting their vision into concrete professional-learning opportunities for teachers; teacher-development programs can prioritize pedagogical content knowledge and the design of mindful, meaningful, and joyful lessons; and supervisors can complement perception surveys with classroom observation. These are suggestions derived from correlational evidence and should be tested with designs that can support causal claims. Several limitations should be considered. (1) The design is cross-sectional, so the study tests associations and predictive contribution rather than improvement over time, and the direction of effects cannot be established. (2) All constructs came from one teacher self-report instrument; common method bias cannot be excluded (full-collinearity VIFs of 3.39 and 3.52), and the discriminant validity of competence versus deep learning quality was not established. (3) Responses showed ceiling effects and limited variance. (4) Teachers were nested in 31 schools with non-trivial intraclass correlations, and although a cluster bootstrap supported the conclusions, a multilevel structural model would be preferable with more schools. (5) School observation scores were not related to teacher-reported scores. (6) The supervisor data came from only three regions and cannot support inferential analysis. Future studies should use multi-source and multi-wave data, separate competence and learning-quality indicators more clearly (for instance, with student or observer ratings of learning), and consider covariance-based or multilevel SEM.

Implication Research

The findings of this study offer several important theoretical implications for the literature on school leadership and instructional quality. Notably, the discovery that teachers' professional competence serves as the primary predictor of deep learning quality (β =0.649) reinforces the premise that mastering pedagogical content and integrating technology effectively are vital for facilitating meaningful educational experiences. Additionally, this study validates the theoretical framework suggesting that a principal's visionary leadership influences student learning indirectly. Rather than directly driving classroom outcomes, visionary leadership operates by shaping a supportive school working environment and fostering the continuous professional development of the teaching staff.

In terms of practical implications, a school leader's vision for deep learning achieves its maximum impact when it is converted into tangible, structured professional growth opportunities for educators. These opportunities should include peer collaboration, constructive clinical supervision, and targeted coaching. Furthermore, teacher training initiatives must prioritize the enhancement of pedagogical content knowledge alongside the practical design of classroom activities that are inherently mindful, meaningful, and joyful.

Limitation Research

Several methodological constraints should be acknowledged when interpreting the findings of this study. First, employing a cross-sectional framework restricts the analysis to evaluating associations and predictive contributions, thereby precluding firm conclusions regarding causal directionality or longitudinal changes over time. Second, because all primary constructs were captured through a single self-reported

instrument, the potential for common method bias cannot be completely ruled out, particularly given that participant responses exhibited noticeable ceiling effects with restricted variance. This measurement constraint is further evidenced by the incomplete discriminant validity between professional competence and deep learning quality, pointing to a probable overlap in how respondents perceived these distinct concepts. Furthermore, because participating teachers were nested within 31 schools featuring non-trivial intraclass correlations, implementing a multilevel structural equation modeling (ML-SEM) strategy would have been methodologically preferable. Finally, the empirical triangulation was constrained by the absence of correlation between school-level observational scores and teacher self-assessments, while data gathered from school supervisors spanned only three supervisory regions, rendering them insufficient for robust inferential statistical analysis.

Future and Suggestion Research

Based on the empirical findings, several practical recommendations are offered for school leaders and policy makers. Principals are strongly encouraged to deliberately align their institutional vision for deep learning with structured, well-planned teacher development programs. Furthermore, educational authorities and school supervisors should move beyond relying solely on perception-based surveys by integrating direct classroom observations and tangible evidence of student-level learning outcomes into their assessment protocols.

To address the inherent constraints of this study, several clear directions for future research are highlighted. Subsequent scholars should adopt longitudinal, multi-source, and multilevel research designs to definitively ascertain the causal sequence among the observed variables. In addition, future investigations must refine the measurement tools for professional competence and deep learning quality to ensure a clearer conceptual separation, potentially by incorporating observer- or student-based evaluations. Finally, upcoming studies should consider employing covariance-based structural equation modeling (CB-SEM) or multilevel SEM frameworks to better account for the nested nature of teacher data distributed across multiple school environments.

CONCLUSION

This study concludes that principals' visionary leadership and teachers' professional competence are significant positive predictors of deep learning quality in elementary schools within Donorojo District. Jointly, these variables account for 71.6% of the variance in deep learning quality. Teachers' professional competence demonstrated a substantial direct effect ($\beta = 0.649$), while visionary leadership yielded a smaller yet significant direct contribution ($\beta = 0.256$), with part of its influence operating indirectly through teacher competence. Although the structural model exhibited good predictive relevance (SRMR = 0.076; $Q^2predict > 0$), certain methodological limitations were identified. Specifically, discriminant validity between professional competence and deep learning quality was not fully established, and school-level observation scores did not entirely corroborate the self-reported data.

Based on these findings, several practical and methodological recommendations are proposed. First, school principals should actively align their vision for deep learning with structured teacher development initiatives, particularly through collaborative learning, instructional supervision, and coaching. Second, teacher training programs must prioritize the enhancement of pedagogical content knowledge and the practical design of mindful, meaningful, and joyful lesson plans. Third, school supervisors and local education offices are advised to evaluate educational quality by triangulating perception surveys with direct classroom observations and student-level evidence. Finally, future research should employ multi-source, multilevel, and longitudinal

designs to definitively confirm causal ordering and to further refine the measurement constructs of professional competence and deep learning quality.

ACKNOWLEDGMENTS

The authors thank the principals, teachers, and school supervisors of Donorojo District, Jepara, for their participation in this study. This research was conducted independently by the researcher and received no specific grant from any funding agency in the public, commercial, or non-profit sectors. This study was conducted in accordance with social science research ethics standards and principles. Written informed consent was obtained from all participants (teachers and school principals) prior to data collection. Participation was voluntary, and the confidentiality of respondent identities and data is fully guaranteed strictly for academic purposes.

REFERENCES

Akmal, A. N., Maelasari, N., & Lusiana, L. (2025). Pemahaman deep learning dalam pendidikan: Analisis literatur melalui metode systematic literature review (SLR). *JIIP - Jurnal Ilmiah Ilmu Pendidikan*, *8*(3), 3229–3236. https://doi.org/10.54371/jiip.v8i3.7442

Al-Fraihat, D., Joy, M., & Sinclair, J. (2020). Evaluating E-learning systems success: An empirical study. *Computers in Human Behavior*, *102*, 67–86. https://doi.org/10.1016/j.chb.2019.08.004

Anwar, M., & Wahyudi, W. (2021). Pengaruh kepemimpinan visioner kepala sekolah terhadap kinerja guru. *Jurnal Administrasi Pendidikan*, *28*(1), 45–56.

Balyer, A. (2012). Transformational leadership behaviors of school principals: A qualitative research based on teachers' perceptions. *International Online Journal of Educational Sciences*, *4*(3), 581–591.

Baumert, J., Kunter, M., Blum, W., Brunner, M., Voss, T., Jordan, A., Klusmann, U., Krauss, S., Neubrand, M., & Tsai, Y.-M. (2010). Teachers' mathematical knowledge, cognitive activation in the classroom, and student progress. *American Educational Research Journal*, *47*(1), 133–180. https://doi.org/10.3102/0002831209345157

Biggs, J., & Tang, C. (2011). *Teaching for quality learning at university* (4th ed.). McGraw-Hill Education.

Bordunos, A. K., Miletich, M. P., & Volkova, N. V. (2024). Mindful learning: Principles and prospect of use in higher education. *Psychological Science and Education*, *29*(4), 16–30. https://doi.org/10.17759/pse.2024290402

Bush, T., & Glover, D. (2014). School leadership models: What do we know? *School Leadership & Management*, *34*(5), 553–571. https://doi.org/10.1080/13632434.2014.928680

Chin, W. W. (2010). How to write up and report PLS analyses. In V. Esposito Vinzi, W. Chin, J. Henseler, & H. Wang (Eds.), *Handbook of partial least squares* (pp. 655–690). Springer. https://doi.org/10.1007/978-3-540-32827-8_29

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.

Darling-Hammond, L. (2017). Teacher education around the world: What can we learn from international practice? *European Journal of Teacher Education*, *40*(3), 291–309. https://doi.org/10.1080/02619768.2017.1315399

Day, C., Gu, Q., & Sammons, P. (2016). The impact of leadership on student outcomes: How successful school leaders use transformational and instructional strategies to make a difference. *Educational Administration Quarterly*, *52*(2), 221–258. https://doi.org/10.1177/0013161X15616863

Dijkstra, T. K., & Henseler, J. (2015). Consistent partial least squares path modeling. *MIS Quarterly*, *39*(2), 297–316. https://doi.org/10.25300/MISQ/2015/39.2.02

Direktorat Jenderal Guru dan Tenaga Kependidikan. (2020). *Pedoman pembinaan dan pengembangan profesi guru*. Kementerian Pendidikan dan Kebudayaan Republik Indonesia.

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, *18*(1), 39–50. https://doi.org/10.1177/002224378101800104

Fullan, M. (2001). *The new meaning of educational change* (3rd ed.). Teachers College Press.

Fullan, M., & Langworthy, M. (2014). *A rich seam: How new pedagogies find deep learning*. Pearson.

Hair, J. F., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2022). *A primer on partial least squares structural equation modeling (PLS-SEM)* (3rd ed.). Sage.

Hair, J. F., Risher, J. J., Sarstedt, M., & Ringle, C. M. (2019). When to use and how to report the results of PLS-SEM. *European Business Review*, *31*(1), 2–24. https://doi.org/10.1108/EBR-11-2018-0203

Hallinger, P. (2011). Leadership for learning: Lessons from 40 years of empirical research. *Journal of Educational Administration*, *49*(2), 125–142. https://doi.org/10.1108/09578231111116699

Hallinger, P., & Heck, R. H. (1998). Exploring the principal's contribution to school effectiveness: 1980–1995. *School Effectiveness and School Improvement*, *9*(2), 157–191. https://doi.org/10.1080/0924345980090203

Hargreaves, A., & Fullan, M. (2012). *Professional capital: Transforming teaching in every school*. Teachers College Press.

Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. Routledge.

Hattie, J., & Donoghue, G. M. (2016). Learning strategies: A synthesis and conceptual model. *npj Science of Learning*, *1*, Article 16013. https://doi.org/10.1038/npjscilearn.2016.13

Hendrianty, B. J., Ibrahim, A., Iskandar, S., & Mulyasari, E. (2024). Membangun pola pikir deep learning guru sekolah dasar. *Kalam Cendekia: Jurnal Ilmiah Pendidikan*, *12*(3), 1348–1358.

Henseler, J., Hubona, G., & Ray, P. A. (2016). Using PLS path modeling in new technology research: Updated guidelines. *Industrial Management & Data Systems*, *116*(1), 2–20. https://doi.org/10.1108/IMDS-09-2015-0382

Henseler, J., Ringle, C. M., & Sarstedt, M. (2015). A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the Academy of Marketing Science*, *43*(1), 115–135. https://doi.org/10.1007/s11747-014-0403-8

Kemendikdasmen. (2025). *Naskah akademik pembelajaran mendalam: Menuju pendidikan bermutu untuk semua*. Pusat Kurikulum dan Pembelajaran, Badan Standar, Kurikulum, dan Asesmen Pendidikan, Kementerian Pendidikan Dasar dan Menengah Republik Indonesia.

Kock, N. (2015). Common method bias in PLS-SEM: A full collinearity assessment approach. *International Journal of e-Collaboration*, *11*(4), 1–10. https://doi.org/10.4018/ijec.2015100101

Kunter, M., Klusmann, U., Baumert, J., Richter, D., Voss, T., & Hachfeld, A. (2013). Professional competence of teachers: Effects on instructional quality and student development. *Journal of Educational Psychology*, *105*(3), 805–820. https://doi.org/10.1037/a0032583

Kusumawati, E. (2023). Analysis of the relationship between the school principal's visionary leadership and kindergarten teachers' performance. *Journal of Innovation in Educational and Cultural Research*, *4*(1), 89–97. https://doi.org/10.46843/jiecr.v4i1.526

Langer, E. J. (2016). *The power of mindful learning* (2nd ed.). Da Capo Press.

Leithwood, K., Harris, A., & Hopkins, D. (2008). Seven strong claims about successful school leadership. *School Leadership & Management*, *28*(1), 27–42. https://doi.org/10.1080/13632430701800060

Leithwood, K., Harris, A., & Hopkins, D. (2020). Seven strong claims about successful school leadership revisited. *School Leadership & Management*, *40*(1), 5–22. https://doi.org/10.1080/13632434.2019.1596077

Lian, B., Sari, A. P., & Rohana. (2025). Pengaruh kepemimpinan visioner dan supervisi akademik terhadap profesionalisme guru SD se-Kecamatan Sako Kota Palembang. *Pendas: Jurnal Ilmiah Pendidikan Dasar*, *10*(2), 229–239. https://doi.org/10.23969/jp.v10i02.28123

Marton, F., & Säljö, R. (1976). On qualitative differences in learning: I—Outcome and process. *British Journal of Educational Psychology*, *46*(1), 4–11. https://doi.org/10.1111/j.2044-8279.1976.tb02980.x

Mehta, J., & Fine, S. (2019). *In search of deeper learning: The quest to remake the American high school*. Harvard University Press.

Mishra, P., & Koehler, M. J. (2006). Technological pedagogical content knowledge: A framework for teacher knowledge. *Teachers College Record*, *108*(6), 1017–1054. https://doi.org/10.1111/j.1467-9620.2006.00684.x

Mulyasa, E. (2013). *Uji kompetensi dan penilaian kinerja guru*. Remaja Rosdakarya.

Mystakidis, S. (2021). Deep meaningful learning. *Encyclopedia*, *1*(3), 988–997. https://doi.org/10.3390/encyclopedia1030075

Nurdin, N., & Sodik, S. (2022). Implementasi deep learning pada kurikulum merdeka di sekolah dasar. *Jurnal Pendidikan Dasar Nusantara*, *8*(2), 112–125.

Ofita, C., & Sururi, S. (2023). Kompetensi pedagogik guru abad 21: Tinjauan peran guru menghadapi generasi alpha. *Jurnal Tata Kelola Pendidikan*, *5*(2), 101–110. https://doi.org/10.17509/jtkp.v5i2.64847

Podsakoff, P. M., MacKenzie, S. B., & Podsakoff, N. P. (2012). Sources of method bias in social science research and recommendations on how to control it. *Annual Review of Psychology*, *63*, 539–569. https://doi.org/10.1146/annurev-psych-120710-100452

Rahmandani, F., Hamzah, M. R., Handayani, T., & Kurniawan, M. W. (2025). Integrasi pembelajaran mendalam (deep learning) dalam mewujudkan pembelajaran yang bermutu dan bermakna bagi peserta didik. *Inovasi: Jurnal Sosial Humaniora dan Pendidikan*, *4*(2), 769–781. https://doi.org/10.55606/inovasi.v4i2.4896

Robinson, V. M. J., Lloyd, C. A., & Rowe, K. J. (2008). The impact of leadership on student outcomes: An analysis of the differential effects of leadership types. *Educational Administration Quarterly*, *44*(5), 635–674. https://doi.org/10.1177/0013161X08321509

Rosyidah, F., Parmiti, D. P., Sanjaya, D. B., & Jampel, I. N. (2026). The theories of principles in deep learning approach: Teori-teori dalam prinsip-prinsip pendekatan pembelajaran mendalam. *Pedagogia: Jurnal Pendidikan*, *15*(1), 127–136. https://doi.org/10.21070/pedagogia.v15i1.2098

Sari, R. P., & Kurniawan, A. (2023). Analisis PLS-SEM untuk mengukur faktor-faktor yang mempengaruhi kompetensi profesional guru. *Jurnal Manajemen Pendidikan*, *11*(1), 34–48.

Shmueli, G., Sarstedt, M., Hair, J. F., Cheah, J.-H., Ting, H., Vaithilingam, S., & Ringle, C. M. (2019). Predictive model assessment in PLS-SEM: Guidelines for using PLSpredict. *European Journal of Marketing*, *53*(11), 2322–2347. https://doi.org/10.1108/EJM-02-2019-0189

Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, *15*(2), 4–14. https://doi.org/10.3102/0013189X015002004

Shulman, L. S. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review*, *57*(1), 1–23. https://doi.org/10.17763/haer.57.1.j463w79r56455411

Supardi, S. (2016). *Kinerja guru*. RajaGrafindo Persada.

Trilling, B., & Fadel, C. (2009). *21st century skills: Learning for life in our times*. John Wiley & Sons.