



Balancing the Load: How Principal Leadership, Welfare, and Workload Shape Teacher Performance in Islamic Primary Schools

Adzkiyaul Maula*, Abu Choir, Adri Efferi, Ashif Az Zafi

Universitas Islam Negeri Sunan Kudus, Indonesia

DOI: <https://doi.org/10.61987/jsse.v4i3.2905>

Article History:

Received: 29 July 2026

Revised: 20 September 2026

Accepted: 16 October 2026

Keywords:

Principal Leadership, Teacher Welfare, Teacher Workload, Teacher Performance

Abstract:

This study examines how madrasah principal leadership, teacher welfare levels, and teacher workload, individually and together, affect teacher performance. This research employed a quantitative approach using a survey research design. The population consisted of all teachers in madrasah, and the study used saturated sampling, selecting all population members as respondents. Data were collected through questionnaires, documentation, and observation. The collected data were analyzed using instrument validity and reliability tests, classical assumption tests, multiple linear regression analysis, t-tests, F-tests, and the coefficient of determination (R^2). The results indicate that madrasah principal leadership has a positive and significant effect on teacher performance. Teacher welfare also significantly influences performance, while workload also has a significant effect: proportional workload can encourage productivity, whereas excessive workload may reduce teaching effectiveness. Simultaneously, madrasah principal leadership, teacher welfare, and workload have a positive and significant effect on teacher performance. The findings imply that improving leadership quality, strengthening teacher welfare, and managing workload distribution effectively are important strategies for achieving sustainable improvement in teacher performance.

INTRODUCTION

Quality education underpins national progress, with teachers as the key figures in shaping that quality (Abdullah & Sukriyah, 2026; Rohmatillah et al., 2026; Ulfa & Hina, 2025). Teachers plan, implement, and evaluate learning processes that directly shape student competence. Empirical evidence indicates that teacher performance relies not only on pedagogical and professional competencies but also on managerial and organizational factors—such as madrasah principal leadership, teacher welfare, and workload (Taufiqurrahman et al., 2025; Widodo, 2024). High-performing teachers are capable of effectively delivering material, guiding students, managing classrooms, developing their own competencies, and responsibly handling administrative tasks. Consequently, understanding the determinants of teacher performance is vital for the broader community, as the findings can serve as a basis for improving graduate quality, enhancing public trust in madrasahs, and fostering sustainable human resource development.

This study employs several theories to explain each variable. Madrasah principal leadership is examined through transformational leadership theory, which encompasses role modeling, inspirational motivation, intellectual stimulation, and individualized consideration (Erfiyana et al., 2026). Welfare is explained through the two-factor theory,

*Corresponding Author:

Adzkiyaul Maula, Universitas Islam Negeri Kudus, Indonesia

Email: adzkiyamaula2@gmail.com

This journal is an open-access article under a CC BY-NC-SA 4.0 license. © 2026, the author(s)

which categorizes compensation, security, and working conditions as hygiene factors and views recognition and personal development as motivators (Ihensekien & Joel, 2023). The job demands–resources model analyzes workload, viewing excessive job demands as triggers for burnout, while supervisor support and job resources serve as buffers (Li et al., 2025). Teacher performance is understood as the work outcomes achieved in fulfilling responsibilities according to established standards. Together, these four frameworks provide a conceptual foundation for examining relationships among the variables in an integrated manner.

Despite the crucial role teachers play, teacher performance varies significantly across many madrasahs—and even among teachers within the same institution. These variations are often linked to differing perceptions regarding the madrasah head's leadership style, the perceived fairness of compensation, allowances, and facilities, and the distribution of workloads across teaching, administrative tasks, and extracurricular activities (Fanani, 2026). Excessive workloads without adequate support lead to burnout and diminished instructional quality, whereas insufficient workloads reduce work effectiveness (Joshi, 2025; Lagawid, 2024). Inadequate welfare also undermines teachers' morale, job satisfaction, and commitment. These conditions create tangible issues for the community, as students receive inconsistent-quality education. Consequently, madrasah administrators require empirical evidence to manage leadership, teacher welfare, and workloads effectively.

Several previous studies have examined the determinants of teacher performance. Efendi & Rifa'i (2025) demonstrated that the transformational leadership of madrasah principals fosters a collaborative work culture, boosts motivation, and encourages teacher professionalism. Hasanuddin et al. (2025) found a significant positive relationship between compensation and welfare and the performance of Madrasah Tsanawiyah teachers. Widiyaningrum et al. (2024) investigated work discipline, the work environment, and job satisfaction among Madrasah Tsanawiyah teachers, while Putra et al. (2024) examined the work environment and workload, with job satisfaction serving as an intervening variable, in the context of Madrasah Aliyah. Anwar et al. (2026) reported that competence, compensation, and work discipline simultaneously influence teacher performance significantly, although work discipline was not significant when analyzed individually. These findings collectively confirm that many factors shape teacher performance.

However, previous research exhibits limitations that need to be addressed. First, most studies have examined variables such as competence, work discipline, and the work environment; consequently, madrasah principal leadership, welfare, and workload have rarely been analyzed within a single model. Second, research on transformational leadership has focused on teacher professionalism rather than its simultaneous impact on performance alongside other organizational factors. Third, the reviewed literature indicates that studies on compensation and workload have predominantly focused on the Madrasah Tsanawiyah and Aliyah levels, while Madrasah Ibtidaiyah has received less attention. These gaps limit the scope of comprehensive policy recommendations. This study addresses these gaps by examining all three factors simultaneously. The novelty of this study lies in the simultaneous examination of three managerial and organizational factors within a single model at the Madrasah Ibtidaiyah (Islamic elementary school) level, while accounting for the interrelationships between these factors. Effective leadership fosters a supportive work environment that helps teachers manage their workloads, while adequate welfare enhances teacher resilience and motivation amid job demands.

This approach moves beyond partial analyses that treat each factor in isolation. Understanding the relative contribution of each variable is crucial for enabling madrasah administrators to prioritize interventions amidst resource constraints accurately. Addressing this issue is urgent, given the persistently high demands regarding the quality of instruction and administration. The findings are expected to enrich the study of Islamic educational management by providing a more comprehensive empirical foundation.

Based on this rationale, the study titled "The Influence of madrasah Principal Leadership, Welfare Levels, and Teacher Workload on Teacher Performance in Madrasah Ibtidaiyah Schools across Karang Tengah District, Demak Regency" was designed to determine whether these three variables influence teacher performance, both individually and simultaneously. The hypotheses propose that madrasah principal leadership and welfare have a positive influence, workload has a significant influence, and all three factors exert a significant simultaneous influence. The underlying argument is that leadership provides direction and support, welfare fosters commitment, and a proportional workload preserves teachers' energy. Theoretically, the study's results are expected to enrich educational management research. Practically, the findings will serve as recommendations for madrasah administrators, policymakers, and initiatives aimed at professional teacher development.

RESEARCH METHODS

This study aims to examine the influence of madrasah principal leadership, welfare levels, and workload on teacher performance (the dependent variable). The study used field research, collecting data directly from the setting where the phenomena occur, specifically teachers' work environment. This approach was selected because teachers' perceptions regarding leadership, welfare, and workload are most accurately obtained directly from the teachers themselves. Consequently, the collected data reflects the actual conditions within the madrasah.

This study employs a quantitative approach with an associative or explanatory design. The quantitative approach emphasizes analyzing numerical data using statistical methods (Almusaed et al., 2025; Oranga & Matere, 2025; Scharrer & Ramasubramanian, 2025). An associative design was used because the study examines cause-and-effect relationships between variables—specifically, the influence of madrasah principal leadership, welfare levels, and workload on teacher performance, both individually (partially) and collectively (simultaneously). The analysis is expected to yield conclusions that are objectively verifiable and defensible.

The study population consists of all Madrasah Ibtidaiyah (Islamic elementary school) teachers in Karang Tengah District, Demak Regency. A population is defined as a general area comprising subjects with specific quantities and characteristics that the researcher uses to study and draw conclusions (Almusaed et al., 2025; Willie, 2024). Based on data regarding active teachers at the time the study was conducted, the population totaled 32 teachers. This number is considered small, and all members were accessible, allowing the researcher to collect data from every teacher without significant obstacles.

The sampling technique employed is a saturated sample (census) method, in which all population members serve as respondents. This technique is appropriate when the population is relatively small—generally ranging from fewer than 30 to 100 individuals—and all members are accessible. Thus, the sample size is equal to the population size—namely, 32 teachers. This technique eliminates sampling error and

improves the representativeness of the data. Consequently, determining the sample size using Slovin's formula is unnecessary.

RESULTS AND DISCUSSION

Results

Validity Testing

We examined item validity using Pearson's product-moment correlation between each item and the scale's total score ($N = 32$). We considered an item valid when its coefficient exceeded the critical r value (.349 for $df = 30$, $\alpha = .05$, two-tailed) and was statistically significant ($p < .05$) (Hair et al., 2019). We also interpreted the coefficients using the proportion of shared variance (r^2). Table 1 presents the results for all four scales.

Table 1. Item-Total Correlations for the Research Instruments (N = 32)

Item	Principal Leadership (X1)	Teacher Welfare (X2)	Workload (X3)	Teacher Performance (Y)
1	.807	.592	.825	.885
2	.881	.771	.886	.859
3	.904	.834	.863	.785
4	.847	.803	.970	.859
5	.811	.854	.858	.874
6	.917	.558	.871	.880
7	.852	.842	.873	.706
8	.857	.822	.933	.809
9	.917	.852	.941	.802
10	.915	.821	.868	.828
<i>M</i>	.871	.775	.889	.829
Range	.807-.917	.558-.854	.825-.970	.706-.885
Valid items	10/10	10/10	10/10	10/10

Note. Critical $r = .349$ ($df = 30$, $\alpha = .05$, two-tailed). All coefficients are significant at $p < .001$, except X2 Item 6 ($p = .001$). *M* = mean of the item-total coefficients.

As Table 1 shows, all 40 items were positively and significantly correlated with their respective scales' total scores, so every item met the validity criteria and was retained. The coefficients ranged from .558 to .970, and the smallest value was still well above the critical value of .349. Therefore, the decision to retain the items does not depend on the chosen significance level.

The workload scale showed the strongest item-total relationships ($M = .889$), followed by principal leadership ($M = .871$), teacher performance ($M = .829$), and teacher welfare ($M = .775$). For workload, three items were exceptionally strong: X3.4 ($r = .970$), X3.9 ($r = .941$), and X3.8 ($r = .933$). Item X3.4 shared 94.1% of its variance with the total score, so it is almost interchangeable with the whole scale. Such values suggest that teachers answered the workload items in a highly consistent pattern. Those who reported a heavy load in one area, such as teaching hours, also tended to report heavy administrative or additional duties. This is plausible because workload components usually co-occur within the same school environment.

The principal leadership items were uniformly strong (.807-.917), and the r^2 values show that each item shared between 65.1% and 84.1% of its variance with the scale. The narrow range indicates that the items contribute evenly to one leadership construct. Teachers' responses to any single item therefore serve as a good proxy for their overall evaluation of the principal's leadership.

The welfare scale had the widest spread of coefficients. Eight items were strong (.771-.854), but X2.6 ($r = .558$, $r^2 = .311$) and X2.1 ($r = .592$, $r^2 = .350$) were noticeably weaker. Both nevertheless exceeded the critical value and were statistically significant. A likely reason is that welfare is multidimensional, covering financial rewards, job security, recognition, professional development, and social relations. An item that taps a single facet will naturally correlate less strongly with a composite of all facets. These two items are the first candidates for review in future refinement, but they were retained here to preserve breadth of content coverage.

The performance items correlated positively with the total score ($r = .706$ to $.885$), with most items sharing 64% to 78% of their variance with the scale. The strongest indicators were Y.1 ($r = .885$), Y.6 ($r = .880$), and Y.5 ($r = .874$). The weakest, Y.7 ($r = .706$, $r^2 = .498$), still showed a strong relationship, about twice the level required for acceptance. This suggests that Y.7 captures a somewhat more specific facet of performance. Because every item remained strongly related to the total score, teachers appear to perceive the different aspects of performance as parts of one overall judgment.

Reliability Testing

Internal consistency was assessed with Cronbach's alpha (Cronbach, 1951). Values of .70 or higher were considered acceptable (Nunnally & Bernstein, 1994), values from .90 to .95 were treated as excellent, and values above .95 were checked for possible item redundancy (Streiner, 2003; Tavakol & Dennick, 2011). Table 2 presents the results.

Table 2. Reliability Coefficients of the Research Instruments (N = 32)

Variable	Number of items	Cronbach's α	Interpretation
Principal leadership (X1)	10	.963	Excellent
Teacher welfare (X2)	10	.925	Excellent
Workload (X3)	10	.969	Excellent
Teacher performance (Y)	10	.948	Excellent

Note. Minimum acceptable $\alpha = .70$.

As shown in Table 2, all four instruments exceeded the minimum threshold of .70 by a wide margin ($\alpha = .925$ -.969). The items within each scale therefore measure the same underlying construct consistently, and the scores would likely be stable if the instruments were administered again under similar conditions.

The welfare scale had the lowest alpha (.925), which is consistent with its validity results. Items X2.1 and X2.6 had the weakest item-total correlations, and a scale covering several welfare facets is expected to be less homogeneous than a unidimensional one. Even so, .925 reflects very high internal consistency. The performance scale ($\alpha = .948$) was also highly reliable, which supports the interpretation that teachers evaluated their performance coherently across its different aspects.

The leadership ($\alpha = .963$) and workload ($\alpha = .969$) scales exceeded .95. Some authors treat such values as a sign that certain items may be redundant, that is, that they ask essentially the same question in different words (Streiner, 2003; Tavakol & Dennick, 2011). This reading fits the very high item-total correlations in Table 1, especially X3.4 ($r = .970$). However, a high alpha is also expected when a scale has a clearly defined construct and ten items. Because the validity evidence shows that the items relate strongly to their constructs, redundancy remains a possibility rather than a demonstrated problem. Future research could shorten these scales without losing reliability. Overall, the four instruments met the criteria for validity and reliability and are suitable for the hypothesis-testing analyses that follow.

Classical Assumption Tests

Before interpreting the regression model, we checked four assumptions: normality of residuals, absence of multicollinearity, homoscedasticity, and independence of residuals. We tested normality with the Kolmogorov-Smirnov test with Lilliefors correction. We assessed multicollinearity using tolerance and the variance inflation factor (VIF), with tolerance $> .10$ and $VIF < 10$ as acceptable criteria. We examined homoscedasticity with the Glejser test, which regresses the absolute residuals on the predictors. Table 3 summarizes the results.

Table 3. Results of the Classical Assumption Tests (N = 32)

Assumption	Test / index	Variable	Value	Criterion	Decision
Normality	Kolmogorov-Smirnov (Lilliefors)	Unstandardized residual	$D = .088, p = .200$	$p > .05$	Met
Multicollinearity	Tolerance / VIF	X1	.838 / 1.193	$Tolerance > .10; VIF < 10$	Met
		X2	.341 / 2.929		Met
		X3	.338 / 2.959		Met
Homoscedasticity	Glejser	X1	$t = 0.856, p = .399$	$p > .05$	Met
		X2	$t = 1.258, p = .219$		Met
		X3	$t = -1.103, p = .279$		Met

Note. X1 = principal leadership; X2 = teacher welfare; X3 = workload. The Lilliefors significance of .200 is a lower bound of the true significance. Dependent variable of the Glejser test = absolute residual.

The normality test did not reject the null hypothesis that the residuals are normally distributed, $D(32) = .088, p = .200$. The residuals had a mean of 0 and a standard deviation of 2.454. The largest positive and negative deviations from the normal distribution were small (.088 and $-.085$), so the residuals can be treated as approximately normal. The t and F tests that follow are therefore appropriate. Because the sample is small, this test has limited power to detect departures from normality, which is why the conclusion is stated as “not rejected” rather than “confirmed.”

All predictors satisfied the multicollinearity criteria. Principal leadership (X1) was nearly independent of the other predictors (tolerance = .838, $VIF = 1.193$), so it contributes largely unique variance to the model. Teacher welfare (X2) and workload (X3) had lower tolerance values (.341 and .338), meaning about 66% of the variance in each is shared with the other predictors. These two variables are therefore moderately to strongly interrelated. Their VIF values (2.929 and 2.959) are still well below the 10 cutoff and below the stricter 5 cutoff. We can thus interpret each coefficient separately, although the standard errors of X2 and X3 are somewhat inflated by their shared variance.

The Glejser test showed no significant relationship between any predictor and the absolute residuals (all $p > .05$; the constant was also non-significant, $p = .778$). Therefore, the error variance does not change systematically with the predictors, and the homoscedasticity assumption holds. The standard errors and t statistics reported below are reliable. Finally, the Durbin-Watson statistic was 2.170 (Table 6). It lies between the upper critical value ($dU = 1.650$ for $n = 32$ and three predictors at $\alpha = .05$) and $4 - dU$ (2.350), so there is no evidence of autocorrelation in the residuals.

Multiple Regression and Partial Hypothesis Tests

We used multiple linear regression to estimate the effects of principal leadership (X1), teacher welfare (X2), and workload (X3) on teacher performance (Y). The partial effects were evaluated with t-tests (two-tailed, $\alpha = .05$, $df = 28$). Table 4 presents the results.

Table 4. Multiple Linear Regression Predicting Teacher Performance (N = 32)

Predictor	<i>B</i>	<i>SE</i>	β	<i>t</i>	<i>p</i>	95% CI for <i>B</i>
(Constant)	5.668	4.433		1.279	.211	[-3.411, 14.747]
X1 Principal leadership	0.271	0.098	.304	2.755	.010	[0.070, 0.472]
X2 Teacher welfare	0.305	0.123	.429	2.482	.019	[0.053, 0.557]
X3 Workload	0.215	0.133	.281	1.616	.117	[-0.057, 0.487]

Note. Dependent variable = teacher performance (Y). *B* = unstandardized coefficient; β = standardized coefficient. Confidence intervals were computed from *B*, *SE*, and the critical *t* value (2.048, $df = 28$).

The estimated regression equation was:

$$Y = 5.668 + 0.271X1 + 0.305X2 + 0.215X3 + e$$

The constant was not significant ($p = .211$). It represents the predicted performance score when all predictors are zero, a value with no practical meaning on these scales and not interpreted further.

Principal leadership (H1). Principal leadership had a significant positive effect on teacher performance, $B = 0.271$, $SE = 0.098$, $\beta = .304$, $t(28) = 2.755$, $p = .010$, 95% CI [0.070, 0.472]. The interval does not include zero. Holding welfare and workload constant, a one-unit increase in perceived leadership was associated with a 0.271-unit increase in teacher performance. H1 was supported. Because X1 had little overlap with the other predictors (tolerance = .838), this effect is largely independent of welfare and workload. It suggests that teachers who perceive their principal as providing direction, motivation, and support tend to report better performance regardless of their welfare or workload.

Teacher welfare (H2). Teacher welfare also had a significant positive effect, $B = 0.305$, $SE = 0.123$, $\beta = .429$, $t(28) = 2.482$, $p = .019$, 95% CI [0.053, 0.557]. Each one-unit increase in welfare was associated with a 0.305-unit increase in performance, holding the other variables constant. H2 was supported. Welfare had the largest standardized coefficient of the three predictors, so within this model it was the strongest predictor of performance. However, the wide confidence interval indicates that the effect size is estimated with limited precision.

Workload (H3). Workload did not have a significant effect, $B = 0.215$, $SE = 0.133$, $\beta = .281$, $t(28) = 1.616$, $p = .117$, 95% CI [-0.057, 0.487]. The interval includes zero, so H3 was not supported at the 5% level. This result should be read with caution. The standardized coefficient (.281) is close to that of leadership (.304), so the non-significance mainly reflects imprecision (a large standard error relative to the coefficient) rather than a clearly negligible effect. Two factors probably contributed: the small sample, and the variance that workload shares with welfare (tolerance = .338). The positive sign indicates that the data provide no evidence that higher workload lowers performance. The data also do not support a positive effect.

In terms of relative contribution, the standardized coefficients ranked welfare (.429), leadership (.304), and workload (.281). Welfare and leadership were the only predictors that remained significant after controlling for the other variables.

Simultaneous Test and Model Fit

The joint effect of the three predictors was tested using an F-test (Table 5), and the model's explanatory power is reported in Table 6.

Table 5. Analysis of Variance for the Regression Model

Source	Sum of squares	df	Mean square	F	p
Regression	466.607	3	155.536	23.337	< .001
Residual	186.612	28	6.665		
Total	653.219	31			

Note. Dependent variable = teacher performance (Y). Predictors: X1, X2, X3.

The regression model was statistically significant, $F(3, 28) = 23.337$, $p < .001$, so H4 was supported. Principal leadership, teacher welfare, and workload jointly predicted teacher performance better than chance. The regression accounted for 466.607 of the 653.219 total sum of squares (71.4%), and the residual mean square was 6.665. The F ratio indicates that the variance explained by each model degree of freedom was more than 23 times the unexplained variance. Note that this joint significance does not depend on workload being significant on its own: the model is meaningful mainly because of leadership and welfare.

Table 6. Model Summary

R	R ²	Adjusted R ²	SE of the estimate	Durbin-Watson
.845	.714	.684	2.582	2.170

Note. Predictors: X1, X2, X3. Dependent variable = teacher performance (Y).

The multiple correlation was strong ($R = .845$). Together, the three predictors explained 71.4% of the variance in teacher performance ($R^2 = .714$), and 68.4% after adjusting for the number of predictors and sample size (adjusted $R^2 = .684$). The small difference between the two values (.030) indicates limited shrinkage, so the number of predictors does not strongly inflate the estimate relative to the sample. This corresponds to a large effect ($f^2 = R^2/(1 - R^2) = 2.50$). The standard error of the estimate (2.582) shows that the model's predictions deviate from observed performance scores by about 2.6 units on average.

The remaining 28.6% of the variance is attributable to factors outside the model. Earlier studies suggest candidates include teacher competence, work discipline, work motivation, and the work environment. In summary, the model fits the data well. Leadership and welfare were significant individual predictors, and workload was not once the other variables were controlled.

Discussion

The model explained 71.4% of the variance in teacher performance, indicating that managerial and organizational conditions account for most of the differences among the teachers sampled. This agrees with Anwar et al. (2026), who found that competence, compensation, and work discipline jointly influenced teacher performance, and with who linked workload and compensation to performance in a madrasah setting. The present study differs in one respect (Firdausiyah et al., 2026; Sarjana et al., 2026). Earlier work mostly relied on individual attributes or tested factors in pairs. In contrast, this model combined leadership, welfare, and workload and showed that leadership and welfare retained unique effects when assessed together. Such a result supports the view that

performance emerges from the interplay of several conditions, not from any single determinant (Achmad et al, 2025; Iltiqoiyah & Thelma, 2025; Meirani & Nurdiana, 2026). Principal leadership had a significant positive effect on performance, consistent with Habeeb & Eyupoglu (2024) and Zadok et al. (2024), who reported that transformational leadership strengthens collaboration, motivation, and professionalism among madrasah teachers. Meta-analytic evidence points the same way: transformational leaders reliably predict follower performance (Alzoraiki et al., 2023; Hafidh Nashrullah & Riyadi, 2026), and school leaders shape instruction mainly by influencing teacher motivation and working conditions (Ilmi, 2024; Rafiqie et al., 2026; Rezeki et al., 2026). The coefficient also aligns with transformational leadership, where role modeling, inspiration, intellectual stimulation, and individualized attention increase commitment beyond formal obligations. Because leadership overlapped little with the other predictors, its influence appears independent of pay or task demands. Variation across studies likely stems from differences in measurement and school context.

Welfare emerged as the strongest predictor, surpassing leadership in standardized terms. This corroborates Alam et al. (2026), both of whom found compensation positively related to teacher performance. Herzberg offer a helpful reading: adequate pay, security, and working conditions prevent dissatisfaction, whereas recognition and development opportunities motivate (Adewojo, 2026; Ebekozien et al., 2026). Because the welfare scale covered both types of elements, its prominence suggests that teachers respond to both material security and symbolic appreciation. Meta-analytic evidence that job satisfaction is moderately related to performance (Liao et al., 2025) further supports this. A caution is warranted, however. The wide confidence interval indicates that the magnitude is estimated imprecisely, so treat the ranking among predictors as tentative rather than definitive.

Workload did not significantly affect performance, which departs who examined workload as a determinant of teacher performance, and from the job demands–resources model, which expects excessive demands to erode performance through exhaustion (Darmadi et al., 2023; Rozi et al., 2025). Several explanations are plausible. Demands differ in kind: challenge demands can stimulate effort, whereas hindrance demands drain it, and a single scale may blend both. Variance shared with welfare, combined with only 32 respondents, also inflated the standard error. The positive coefficient suggests that teachers may perceive tasks as meaningful responsibilities when leaders and rewards support them. Nonetheless, a non-significant result does not prove no effect, so replication with larger samples is needed.

These findings make three theoretical contributions. First, they extend transformational leadership theory to madrasah ibtidaiyah, a primary-level Islamic school context that remains underrepresented in the literature. Second, they show that Herzberg’s two-factor logic and the job demands–resources framework can be tested in a single equation, revealing that resources (leadership and welfare) mattered more than demands (workload) in this sample. Third, the divergent workload result refines the framework by suggesting that its predictions depend on how demands are appraised and on the resources available to meet them. Researchers should therefore separate challenge from hindrance demands in future instruments, and consider mediators such as job satisfaction or engagement, as well as moderating roles for leadership, to capture mechanisms that a direct-effects model cannot.

Practically, the results give madrasah administrators a clear order of priorities. Because welfare carried the largest coefficient, foundations and ministry offices should review allowances, incentive schemes, job security, and recognition programs, since these appear to pay off in performance. Principals can strengthen their influence through mentoring, participative decision-making, and individualized feedback, which turn leadership into everyday support. Workload need not be the first target for intervention, though fair task distribution remains sound management. Several limitations qualify these recommendations: the cross-sectional design, self-reported perceptions, and a sample of 32 teachers restrict causal claims and generalization. Future studies could combine larger samples, supervisor ratings, and longitudinal designs to test whether the observed patterns endure.

CONCLUSION

This study shows that teacher performance in madrasah depends more on the support teachers receive than on the demands placed on them. Principal leadership and teacher welfare each had a significant positive effect on performance, with welfare the strongest predictor, while workload had no significant effect; together, the three variables explained 71.4% of the variance. The central lesson is that teachers perform best when leaders guide them and when their material and symbolic needs are met. Scientifically, the study contributes an integrated model that tests transformational leadership, two-factor motivation, and job demands–resources logic in one equation within the under-researched context of madrasah ibtidaiyah. Practically, foundations and ministry offices should prioritize allowances, incentives, job security, and recognition, while principals should adopt mentoring, participative decision-making, and individualized feedback. Several limitations qualify these conclusions: the cross-sectional design, reliance on self-reported perceptions, and a sample of only 32 teachers restrict causal inference and generalization, and the non-significant workload result may reflect low statistical power. Future research should use larger, more diverse samples; supervisor-rated or objective performance data; and longitudinal designs, while separating challenge from hindrance demands and testing mediators such as job satisfaction and engagement, to establish the mechanisms behind these relationships more firmly.

ACKNOWLEDGEMENT

The authors thank Universitas Islam Negeri (UIN) Sunan Kudus for its support in preparing this study. The authors also thank the teachers who took part as respondents for their time and cooperation.

CONFLICTS OF INTEREST

No conflict of interest.

USE OF ARTIFICIAL INTELLIGENCE (AI)-ASSISTED TECHNOLOGY

The authors utilized Claude (Anthropic) and Grammarly as AI-assisted tools to enhance language quality, correct grammatical errors, and improve the clarity, wording, and organization of the manuscript. All AI-generated suggestions were carefully reviewed and revised by the authors, who assume full responsibility for the accuracy, originality, and integrity of the published content.

REFERENCES

- Abdullah, A., & Sukriyah, T. M. (2026). Soft Skill Management in Improving the Quality and Competence of Human Resources in Islamic Junior High Schools. *Business and Applied Management Journal*, 4(1), 134–144. <https://doi.org/10.61987/bamj.v4i1.2266>
- Achmad, F., Riau, D. P., & Akadira, T. (2025). Evaluating Public Service Quality: Why Employee Performance Outshines Service Standards and Work Culture in Shaping User Satisfaction. *Business and Applied Management Journal*, 3(2), 349–362. <https://doi.org/10.61987/bamj.v3i2.1638>
- Adewojo, A. (2026). Working Environment and Job Satisfaction of Library Personnel in Research Institutes in South-West Nigeria. *International Information & Library Review*, 1–18. <https://doi.org/10.1080/10572317.2026.2710535>
- Alam, A. R. N., Agustina, R., & Sudirman, M. A. (2026). Compensation as a Determinant of Teacher Performance: A Case Study in Public Education. *GENTARA: Management and Accounting Research*, 1(1), 61–67.
- Almusaed, A., Almssad, A., & Yitmen, I. (2025). Introduction to Quantitative Research. In *Practice of Research Methodology in Civil Engineering and Architecture: A Comprehensive Guide* (pp. 489–528). Springer. https://doi.org/10.1007/978-3-031-97393-2_15
- Alzoraiki, M., Ahmad, A. R., Ateeq, A. A., Naji, G. M. A., Almaamari, Q., & Beshr, B. A. H. (2023). Impact of Teachers' Commitment to the Relationship Between Transformational Leadership and Sustainable Teaching Performance. *Sustainability*, 15(5), Article 4620. <https://doi.org/10.3390/su15054620>
- Anwar, K., Syakur, A., Mahmudah, I., & Priyono, A. A. (2026). Structural Model of Teacher Work Productivity Based on Discipline and Compensation. *Nidhomul Haq: Jurnal Manajemen Pendidikan Islam*, 11(1), 163–175. <https://doi.org/10.31538/ndhq.v11i1.375>
- Darmadi, D., Pratama, G. D., Susanto, N., & Fitria, J. R. (2023). The Impact of Transformational Leadership Style and Work Environment on Employee Performance. *Journal of Economics and Business Letters*, 3(5), 338–348. <https://doi.org/10.55942/jebli.v3i5.241>
- Ebekozien, A., Hafez, M., Aigbavboa, C., Ebekozien, E. O., Ghazali, N., Mapfumo, E. T., Nwaole, A. N. C., Samsurijan, M. S., Eremiokhale, J. I., & Aigbe, R. O. (2026). Appraising Artisans' Job Satisfaction in Infrastructure Provisions to Achieve SDGs 3 and 8: The Role of Herzberg's Motivation-Hygiene Theory. *Engineering, Construction and Architectural Management*, 1–19. <https://doi.org/10.1108/ECAM-11-2025-1764>
- Efendi, D. R., & Rifa'i, A. A. (2025). Transformational Leadership of School Principals in the Development of Islamic Education. *Nidhomiyah: Jurnal Manajemen Pendidikan Islam*, 6(2), 240–252. <https://doi.org/10.38073/nidhomiyah.v6i2.2719>
- Erfiyana, E., Hoerudin, C. W., Arifudin, O., & Kartika, I. (2026). Transformational Leadership of School Principals in Developing Islamic Education in Elementary Madrasahs. *International Journal of Science Education and Technology Management (IJSETM)*, 5(1), 32–44.
- Fanani, S. K. (2026). Managerial Strategies in Teacher Welfare Policies for Islamic Education: A Comparative Study of Indonesia and Malaysia. *Benchmarking*, 10(1), 130–137. <https://doi.org/10.30821/benchmarking.v10i1.28348>

- Firdausiyah, J., Rifka, Khomairohtusshiyamah, U., & Sain, Z. H. (2026). Strategic Management in Developing Inclusive Teacher Career Paths: Optimizing Potential and Performance in Educational Institutions. *Journal of Psychological Insight*, 2(1), 39–50. <https://doi.org/10.61987/jpi.v2i1.1477>
- Habeeb, Y. O., & Eyupoglu, S. Z. (2024). Strategic Planning, Transformational Leadership, and Organization Performance: Driving Forces for Sustainability in Higher Education in Nigeria. *Sustainability*, 16(11), Article 4348. <https://doi.org/10.3390/su16114348>
- Hafidh Nashrullah, M., & Riyadi, S. (2026). From Transformational Leadership to Social Entrepreneurship: Strengthening Innovation and Sustainability in Faith-Based Organizations. *Business and Applied Management Journal*, 4(1), 16–32. <https://doi.org/10.61987/bamj.v4i1.2041>
- Hasanuddin, F. N. M., Amin, A. M., Natsir, U. D., Idris, A. A., & Hamka, R. A. (2025). The Influence of Financial and Non-Financial Compensation on Teacher Performance at Madrasah Tsanawiyah Negeri 1 Muna. *International Journal of Economic Literature*, 2(7), 741–749.
- Ihensekien, O. A., & Joel, A. C. (2023). Abraham Maslow's Hierarchy of Needs and Frederick Herzberg's Two-Factor Motivation Theories: Implications for Organizational Performance. *The Romanian Economic Journal*, 85(1), 133–146. <https://doi.org/10.24818/REJ/2023/85/04>
- Ilmi, K. (2024). Interpersonal Communication of Madrasah Heads in Increasing Teacher Work Motivation. *Business and Applied Management Journal*, 1(2), 104–117. <https://doi.org/10.61987/bamj.v1i2.480>
- Ilitqoiyah, L., & Thelma, C. C. (2025). The Relationship Between Technical Skills of Madrasah Heads and Non-Financial Compensation with the Performance of Private Islamic Elementary School Teachers. *Journal of Social Studies and Education*, 3(1), 58–70. <https://doi.org/10.61987/jsse.v3i1.677>
- Joshi, R. P. (2025). Exploring Teacher Burnout: A Comprehensive Study of Causes and Solutions. *EDUCASIA: Jurnal Pendidikan, Pengajaran, dan Pembelajaran*, 10(2), 245–262. <https://doi.org/10.21462/educasia.v10i2.367>
- Lagawid, E. L. (2024). Overloading in Elementary School Teachers: A Review. *Cognizance Journal of Multidisciplinary Studies*, 4(11), 289–299. <https://doi.org/10.47760/cognizance.2024.v04i11.016>
- Li, Y., Chen, C., & Yuan, Y. (2025). Evolving the Job Demands-Resources Framework to JD-R 3.0: The Impact of After-Hours Connectivity and Organizational Support on Employee Psychological Distress. *Acta Psychologica*, 253, Article 104710. <https://doi.org/10.1016/j.actpsy.2025.104710>
- Liao, Y.-K., Wang, S., Thinh, V. T., & Wu, W.-Y. (2025). An Investigation of Leading and Lagging Indicators for Balanced Scorecard: A Meta-Analytic Approach. *SAGE Open*, 15(4), Article 21582440251359830. <https://doi.org/10.1177/21582440251359833>
- Meirani, R. A., & Nurdiana, I. I. (2026). Beyond Financial Management: An Integrated Framework for Strengthening Financial Performance in Educational Institutions. *Journal of Social Studies and Education*, 4(1), 167–179. <https://doi.org/10.61987/jsse.v4i1.2967>
- Oranga, J., & Matere, A. (2025). Quantitative Research: Types, Advantages, Generalizability, and Limitations. *Open Access Library Journal*, 12(9), 1–9. <https://doi.org/10.4236/oalib.1113694>

- Putra, R. S., Eliyana, A., Agustina, T. S., Khusnah, H., & Anshori, M. Y. (2024). The Influence of Workload and Working Environment on Employee Performance Through Job Satisfaction as a Mediation Variable at CV. Kebab Bosman Food Indonesia. *Revista de Gestão Social e Ambiental*, 18(5), Article e05633. <https://doi.org/10.24857/rgsa.v18n5-088>
- Rafiqie, M., Haryanti, N., Ihwan, M., Habibi, I., & Sutarum, S. (2026). Beyond Technicalities: Spiritual Motivation as a Mediator Between Academic Supervision and Teacher Performance. *Al-Tanzim: Jurnal Manajemen Pendidikan Islam*, 10(3), 906–918. <https://doi.org/10.33650/al-tanzim.v10i3.14318>
- Rezeki, D. P., Mesiono, M., & Khadijah, K. (2026). The Influence of Organizational Culture, Principal Leadership, Digital Literacy, and Work Motivation on Teacher Integrity. *Al-Tanzim: Jurnal Manajemen Pendidikan Islam*, 10(3), 919–934. <https://doi.org/10.33650/al-tanzim.v10i3.13793>
- Rohmatillah, L. F., Adeoye, M. A., & Iliasu, S. F. (2026). Beyond Administration: How Continuous Quality Improvement Transforms Professional Collaboration and Student Engagement in Elementary Schools. *Al-Tanzim: Jurnal Manajemen Pendidikan Islam*, 10(2), 479–493. <https://doi.org/10.33650/al-tanzim.v10i2.13658>
- Rozi, F., Hina, S., & Paramole, O. C. (2025). Overcoming Teacher Burnout to Improve Performance Through Spiritual Approach Services. *Business and Applied Management Journal*, 2(2), 60–70. <https://doi.org/10.61987/bamj.v2i2.504>
- Sarjana, B., Sholihah, F. N., & Meishanti, O. P. Y. (2026). Fostering Student-Centered Math Classrooms: The Role of Kahoot in Optimizing Academic Performance. *Indonesian Journal of Progressive Pedagogy*, 2(2), 127–138. <https://doi.org/10.61987/ijpp.v2i2.1757>
- Scharrer, E., & Ramasubramanian, S. (2025). *Quantitative Research Methods in Communication: The Power of Numbers for Social Justice*. Routledge. <https://doi.org/10.4324/9781003459453>
- Taufiqurrahman, M., Ridani, A., & Sudadi, S. (2025). Exploring the Impact of Madrasah Head's Leadership Style on Teacher Performance. *Indonesian Journal of Education and Social Studies*, 4(2), 180–193. <https://doi.org/10.33650/ijess.v4i2.12977>
- Ulfa, M., & Hina, S. (2025). The Role of Artificial Intelligence (AI) in Improving the Quality of Education in Early Childhood Education Programs. *Indonesian Journal of Progressive Pedagogy*, 1(1), 49–59. <https://doi.org/10.61987/ijpp.v1i1.666>
- Widiyaningrum, A., Kholiq, A., & Pratiwi, D. (2024). The Influence of Professional Competence, Workload, and Work Commitment on Teacher Performance Through Job Satisfaction in Madrasah Tsanawiyah. *Abjadia: International Journal of Education*, 9(2), 433–448. <https://doi.org/10.18860/abj.v9i2.28105>
- Widodo, W. (2024). Developing Teacher Contextual Performance Through School Organizational Culture and Cultural Intelligence. *Al-Tanzim: Jurnal Manajemen Pendidikan Islam*, 8(1), 59–72. <https://doi.org/10.33650/al-tanzim.v8i1.5811>
- Willie, M. M. (2024). Population and Target Population in Research Methodology. *Golden Ratio of Social Science and Education*, 4(1), 75–79. <https://doi.org/10.52970/grsse.v4i1.405>
- Zadok, A., Benoliel, P., & Schechter, C. (2024). School Middle Leaders' Transformational Leadership and Organizational Resilience: The Moderating Role of Academic Emphasis. *European Journal of Education*, 59(3), Article e12657. <https://doi.org/10.1111/ejed.12657>