Conceptual Article

EDUCATIONAL POINT
An International Peer-Reviewed Journal

https://www.educationalpoint.net
ISSN: 3062-1011 (Online)

# Trust as a Mediating Mechanism in AI-Enabled School Leadership: Navigating Benefits, Risks, and Ethical Tensions in Education 4.0

Nicos Keravnos [1*], Maria Eleftheriou [1]

[1] School of Education, Frederick University, CYPRUS
* Correspondence: nicoskeravnos@gmail.com

**CITATION:** Keravnos, N., & Eleftheriou, M. (2026). Trust as a Mediating Mechanism in AI-Enabled School Leadership: Navigating Benefits, Risks, and Ethical Tensions in Education 4.0. *Educational Point*, *3*(3), e182. https://doi.org/10.71176/edup/18914

**ARTICLE INFO**

Received: 26 March 2026
Accepted: 27 April 2026

**OPEN ACCESS**

**ABSTRACT**

Artificial intelligence (AI) is rapidly reshaping school leadership within Education 4.0, offering enhanced decision-making and organisational efficiency while intensifying ethical concerns regarding transparency, bias, and accountability. Existing research has largely treated these opportunities and risks as separate phenomena, overlooking the relational processes through which AI is enacted in practice. This paper advances a process-based conceptualisation by positioning trust as the central mediating mechanism in AI-enabled school leadership. It argues that AI does not produce outcomes directly; rather, its effects are contingent on how it is accepted, interpreted, and enacted within school contexts. The proposed framework shows that trust shapes whether AI leads to constructive outcomes, including ethical use, professional engagement, and improvement, or to disruptive consequences such as resistance and mistrust. Leadership is conceptualised as a key antecedent of trust, highlighting the centrality of relational governance in the effective and responsible integration of AI in schools.

**Keywords:** artificial intelligence, school leadership, trust, educational leadership, ai governance, education 4.0

# INTRODUCTION

Artificial intelligence (AI) is rapidly reshaping school leadership within the broader context of Education 4.0, offering new opportunities for data-informed decision-making, organisational efficiency, and strategic planning. As digital technologies become increasingly embedded in educational systems, school leaders are expected to navigate complex data environments, utilise predictive analytics, and coordinate more responsive organisational processes. Emerging research suggests that AI-enabled systems can support leaders in monitoring performance, optimising resource allocation, and enhancing administrative decision-making, thereby strengthening evidence-informed leadership practices (Adams & Thompson, 2025; Dai et al., 2025; Kesim et al., 2025; Sain et al., 2024). These developments align with broader conceptualisations of leadership

as the work of setting direction, developing people, and redesigning organisational processes. Together, these functions aim to improve learning outcomes (Leithwood et al., 2004).

However, the integration of AI into school leadership extends beyond technical enhancement and introduces significant organisational and ethical challenges. Existing research highlights concerns related to data privacy, algorithmic bias, opacity in automated decision-making, and the diffusion of accountability across human and technological actors (Fu & Weng, 2024; García-López & Trujillo-Liñán, 2025; Wang, 2024; Williamson et al., 2024). These challenges are not purely technical or regulatory; they reshape professional relationships and influence how educators interpret and respond to leadership decisions involving AI. AI adoption, therefore, represents a socio-organisational transformation that redefines how decisions are produced, justified, and enacted within schools (Arar et al., 2025; Fullan et al., 2024).

Despite growing interest in AI in education, existing scholarship has largely examined its opportunities and risks in isolation. Studies often emphasise either efficiency and innovation or ethical and governance concerns, without adequately addressing the relational dynamics that shape how AI is enacted in practice. As a result, limited attention has been given to the mechanisms through which educators interpret AI, negotiate its uncertainties, and decide whether to integrate it into their professional work. This represents a critical gap, as technological initiatives in schools are mediated by organisational relationships and professional cultures rather than implemented in a vacuum.

Trust, therefore, provides a useful lens for addressing this gap. It has long been recognised as a foundational resource for effective school leadership and improvement, supporting collaboration, professional engagement, and collective problem-solving (Bryk & Schneider, 2002; Tschannen-Moran & Gareis, 2015). In high-trust environments, educators are more willing to take risks, engage in innovation, and participate in organisational change processes. Trust is not a static attribute but is actively constructed through leadership practices and everyday interactions, grounded in perceptions of competence, integrity, and benevolence (Mayer et al., 1995), as well as fairness, transparency, and professional support (Tschannen-Moran, 2014).

In AI-enabled contexts, trust becomes even more critical. Educators must evaluate not only the technical reliability of AI systems but also the credibility, fairness, and intentions of the leadership guiding their implementation. The inherent uncertainty and ethical ambiguity associated with AI heighten the importance of trust as a mechanism through which educators make sense of and engage with technological change.

Building on this perspective, this paper argues that trust functions as a mediating mechanism between the potential benefits and inherent risks of AI in school leadership. Rather than producing outcomes directly, AI is enacted through relational processes that shape how it is accepted, interpreted, and used within school contexts. In this view, the impact of AI is not technologically predetermined but socially constructed.

This study makes three key contributions. First, it moves beyond technocentric and risk-based accounts by introducing a relational, process-oriented perspective on AI implementation. Second, it reconceptualises trust as an active mediating mechanism operating through the distinct processes of acceptance, interpretation, and enactment, offering greater analytical precision than existing frameworks. Third, it integrates insights from AI, trust, and leadership literatures into a unified conceptual model, shifting the field from descriptive accounts of AI adoption toward an organisational explanation of how AI is enacted in practice. Collectively, these contributions position trust at the centre of AI-enabled school leadership and highlight the relational conditions necessary for responsible and effective technological integration.

# LITERATURE REVIEW: ARTIFICIAL INTELLIGENCE AND TRUST IN SCHOOL LEADERSHIP

This section reviews and integrates literature on artificial intelligence (AI) in education and trust in school leadership to establish the conceptual foundations of the study. Rather than treating these domains as separate, the review adopts a relational perspective, examining how technological developments and organisational dynamics intersect in school contexts. It first explores the opportunities and organisational implications of AI in school leadership, then examines the ethical and governance challenges associated with its use. The review subsequently considers trust as a central relational construct in educational leadership before synthesising these strands to identify a conceptual gap and justify the development of the proposed framework.

## Artificial Intelligence in School Leadership: Opportunities and Transformation

Artificial intelligence (AI) has become an increasingly prominent feature of Education 4.0, reshaping expectations of school leadership and organisational practice. Within this evolving landscape, AI technologies are frequently positioned as tools that enhance leaders' capacity to manage complexity through data-informed decision-making, predictive analytics, and administrative optimisation. Empirical studies indicate that AI can support school leaders in identifying trends in student performance, allocating resources more strategically, and streamlining organisational processes (Adams & Thompson, 2025; Dai et al., 2025; Sain et al., 2024). These developments reflect a broader shift toward evidence-informed leadership, in which decisions are increasingly supported by real-time data and advanced analytics.

However, the assumption that analytical capacity automatically translates into improved organisational outcomes remains contested. While some studies emphasise AI's potential to enhance rational decision-making, others caution that increased reliance on data-driven systems may marginalise professional judgement. This tension highlights a fundamental challenge between efficiency and professional autonomy in AI-enabled environments.

Beyond administrative functions, AI also influences broader organisational dynamics. Research suggests that AI adoption can reshape communication patterns, redefine professional roles, and alter leadership practices (Kesim et al., 2025; Lipsou et al., 2026). These effects position AI not simply as a technical innovation but as an organisational force that interacts with culture, professional norms, and relational processes. Accordingly, the impact of AI cannot be understood solely through its technical capabilities. Its influence emerges through professional interpretation and organisational context.

## Ethical Challenges and Governance Tensions in AI-Enabled Education

Alongside its potential benefits, AI introduces significant ethical and governance challenges. A growing body of literature highlights concerns related to algorithmic bias, data privacy, opacity in automated decision-making, and the diffusion of accountability across human and technological actors (Fu & Weng, 2024; García-López & Trujillo-Liñán, 2025; Wang, 2024). These challenges are particularly salient when AI systems function as "black boxes," limiting educators' ability to understand or critically evaluate decision outputs.

The governance of AI further complicates implementation. Responsibility for AI-assisted decisions is often distributed across multiple actors, including system designers, policymakers, and school leaders. This diffusion of accountability challenges traditional governance structures and places new demands on leaders to ensure ethical and responsible use of AI (Williamson et al., 2024). At the same time, AI-driven monitoring

systems raise concerns about surveillance, autonomy, and professional agency, potentially undermining the relational foundations of educational practice.

While existing scholarship provides valuable insights into these ethical and structural dimensions, it tends to conceptualise AI implementation in technical or regulatory terms. Less attention has been given to how educators interpret these challenges in everyday practice, how they negotiate uncertainty, and how organisational relationships shape responses to AI. This limitation underscores the need for perspectives that integrate ethical, organisational, and relational dimensions.

## Trust in School Leadership: A Relational Foundation for Change

Trust has long been recognised as a central resource for effective school leadership and organisational improvement (Louis, 2007). Foundational research demonstrates that relational trust supports collaboration, professional engagement, and sustained school improvement (Bryk & Schneider, 2002). Subsequent studies confirm that trust enhances teacher commitment, facilitates professional learning, and supports openness to innovation (Louis & Murphy, 2017; Tschannen-Moran & Gareis, 2015). In high-trust environments, educators are more likely to take risks, share expertise, and engage in collective problem-solving (Bukko et al., 2021).

Trust is actively constructed through leadership practices and everyday interactions. It is grounded in perceptions of competence, integrity, and benevolence (Mayer et al., 1995), as well as in behaviours such as fairness, transparency, and professional support (Tschannen-Moran, 2014). Recent research further highlights how trust is reinforced through consistent communication, responsiveness, and inclusive decision-making processes (Keravnos & Symeou, 2026). Leadership approaches that integrate transformational, instructional, and distributed practices have been shown to strengthen trust by enhancing participation, professional credibility, and shared ownership (Keravnos et al., 2025).

Despite its recognised importance, trust is often conceptualised as a general enabling condition rather than an active process. Limited attention has been given to how trust operates within the implementation of innovation—how it shapes sensemaking, influences engagement, and mitigates uncertainty. This limitation becomes particularly significant in AI-enabled contexts, where uncertainty is heightened and professional judgement is challenged.

## Trust and Artificial Intelligence: Bridging a Conceptual Divide

When considered together, the literature on AI in education and trust in leadership reveal a notable disconnect. AI research tends to focus on technological capabilities and governance challenges, whereas trust research emphasises relational dynamics and organisational processes. Rarely are these perspectives integrated. As a result, existing frameworks do not fully explain how AI is interpreted, negotiated, and enacted within school contexts.

AI introduces new forms of epistemic and ethical uncertainty. The opacity of algorithmic processes, the potential for biased outcomes, and concerns about the erosion of professional judgement create conditions in which trust becomes both more fragile and more necessary (Wang, 2024; Williamson et al., 2024). Under these conditions, educators must determine not only whether they trust the technology but also whether they trust the leadership decisions surrounding its implementation.

Emerging research suggests that trust shapes how uncertainty is experienced and how innovation is enacted in organisations. High levels of trust can enable educators to interpret uncertainty as an opportunity for learning and improvement, whereas low trust may lead to scepticism, resistance, or disengagement. These dynamics are particularly salient in AI-enabled environments, where meaning is not inherent in the technology but constructed through organisational relationships.

This synthesis highlights the need for a conceptual framework that explains how trust mediates the relationship between AI and organisational outcomes. Rather than treating AI adoption as a linear or purely technical process, such a framework must account for the relational mechanisms through which educators accept, interpret, and enact AI systems. This perspective shifts the analytical focus from technological capability to relational mediation, providing a more nuanced understanding of how AI becomes embedded in leadership practice.

# CONCEPTUAL FRAMEWORK: TRUST AS A MEDIATING MECHANISM IN AI-ENABLED SCHOOL LEADERSHIP

This section presents a conceptual framework that positions trust as the central mechanism shaping how artificial intelligence (AI) is enacted in school leadership. Building on research in AI in education, educational leadership, and organisational trust, the framework advances a process-based explanation of how AI's opportunities and risks are translated into school-level practice. Rather than assuming that AI produces direct or technologically determined outcomes, the framework conceptualises AI implementation as a socially constructed process shaped by leadership practices, trust dynamics, and organisational sensemaking. This approach addresses a critical gap in existing scholarship, which has tended to treat AI as either a technical enhancement or an ethical problem, without examining the relational processes through which educators navigate AI's inherent dualities. To address this, the framework integrates leadership practices, multidimensional trust, and three mediating processes, acceptance, interpretation, and enactment, offering a coherent and testable model for understanding AI implementation in schools.

## Framework Development and Analytical Approach

The framework was developed through an integrative conceptual approach (Torraco, 2005, 2016). Literature was identified through structured searches in Scopus, Web of Science, and Google Scholar using Boolean combinations of key terms, including "artificial intelligence AND education," "AI AND school leadership," "trust AND educational leadership," and "AI governance AND schools." The review included peer-reviewed English-language publications from 2002 to 2026 and incorporated seminal theoretical contributions where conceptually necessary. Core sources included foundational trust theories (e.g., Bryk & Schneider, 2002; Mayer et al., 1995; Tschannen-Moran, 2014), recent work on AI-enabled school leadership (e.g., Adams & Thompson, 2025; Dai et al., 2025; Fullan et al., 2024; Kesim et al., 2025; Lipsou et al., 2026; Williamson et al., 2024), and research examining trust in AI technologies (e.g., Choung et al., 2023; Shin, 2021). Additional insights were drawn from studies on relational trust (e.g., Connelly et al., 2018; Cranston, 2011; Edwards-Groves et al., 2020), leadership integration (e.g., Keravnos et al., 2025; Marks & Printy, 2003), and distributed accountability in technologised governance environments (e.g., Fu & Weng, 2024; Wang, 2024). Non-education-focused studies and duplicate records were excluded. The selection process emphasised conceptual relevance, theoretical contribution, and recency, ensuring a balanced integration of foundational and emerging scholarship.

The synthesis proceeded iteratively. Mapping the documented opportunities and risks of AI against the well-established role of trust in organisational change revealed a notable absence of process-based explanations in existing literature. To address this gap, the mediating processes of acceptance, interpretation, and enactment were developed as analytically distinct mechanisms that explain how trust shapes educators' engagement with AI. By identifying these processes, the framework moves beyond descriptive summaries of AI implementation and articulates causal pathways that can support empirical testing. The focus is bound to school-level (K–12) implementation in Education 4.0 contexts characterised by moderate AI maturity, positioning the framework as a theory-building contribution intended to guide future empirical work.

## From Technological Duality to Relational Mediation

AI introduces a well-documented duality into school leadership. On one side, AI offers enhanced decision-making, organisational efficiency, and new forms of data-informed practice (Adams & Thompson, 2025; Dai et al., 2025; Sain et al., 2024). On the other hand, AI raises concerns about privacy, algorithmic bias, opacity, and the diffuse and often ambiguous distribution of accountability among leaders, systems, and external providers (Fu & Weng, 2024; García-López & Trujillo-Liñán, 2025; Wang, 2024; Williamson et al., 2024). Rather than treating these dimensions as independent, the framework conceptualises them as a single condition of uncertainty in which benefits and risks are intertwined.

Under such conditions, implementation outcomes are not determined by the technology itself but emerge through organisational relationships and interpretive practices. Existing research often privileges either technical capability or ethical critique, thereby overlooking the relational processes through which educators make sense of AI. In school contexts, where leadership is inherently relational, collaborative, and trust-dependent, these relational processes are central to understanding how AI is enacted (Yılmaz, 2025). The framework, therefore, emphasises relational mediation as the mechanism through which AI's duality is navigated.

## Trust as a Relational Governance Mechanism

Trust is conceptualised as a key form of relational governance that supports coordinated action under uncertainty. Grounded in perceptions of competence, integrity, and benevolence (Mayer et al., 1995), trust is essential in contexts where formal control is limited (Bryk & Schneider, 2002; Handford & Leithwood, 2013; Tschannen-Moran, 2014). In schools, trust develops through repeated exchanges characterised by respect, fairness, professionalism, and care (Bryk & Schneider, 2002; Keravnos & Symeou, 2026). Leadership practices play a central role in cultivating trust: integrated approaches that combine transformational, instructional, and distributed leadership have been shown to strengthen participation, transparency, and professional credibility (Keravnos et al., 2025; Marks & Printy, 2003).

Trust also functions as an informal governance mechanism. High-trust environments foster engagement, collaboration, responsible risk-taking, and openness to innovation (Ellonen et al., 2008; Tschannen-Moran & Gareis, 2015), reducing reliance on formal monitoring and enabling collective problem-solving. In AI-enabled contexts, this governance function becomes intensified. AI's opacity and complexity introduce uncertainty about how decisions are generated and who is accountable (Wang, 2024; Williamson et al., 2024), making trust essential for educators to meaningfully engage with AI outputs.

Trust is multidimensional. Trust in leadership reflects confidence in leaders' intentions, competence, and ethical conduct (Bligh, 2017; Caldwell et al., 2008; Kutsyuruba & Walker, 2015). Trust in technology reflects perceptions of reliability, fairness, explainability, and transparency (Choung et al., 2023; Lipsou et al., 2026; Shin, 2021). Relational trust refers to shared norms of respect, collaboration, and integrity within the professional community (Connelly et al., 2018; Cranston, 2011; Edwards-Groves et al., 2020). These dimensions interact to shape how educators interpret AI initiatives and decide whether to engage with them.

## The Mediating Function of Trust in AI Implementation

Integrated leadership practices shape the development of multidimensional trust, which then mediates three interrelated processes through which AI is enacted: acceptance, interpretation, and enactment. Acceptance concerns educators' initial willingness to engage with AI systems. Trust in leadership and relational trust influence whether AI is perceived as legitimate, aligned with professional values, and introduced through fair and transparent processes (Tschannen-Moran & Gareis, 2015; Williamson et al., 2024).

Interpretation refers to the sensemaking processes through which educators evaluate AI outputs. Trust in technology shapes perceptions of credibility, reliability, and fairness (Fu & Weng, 2024; Shin, 2021), while relational trust supports collaborative reasoning and the integration of algorithmic insights with contextual knowledge and professional judgement (Griffiths et al., 2021; Wang, 2024). Enactment concerns the practical use of AI in daily organisational routines. High-trust environments enable context-sensitive, ethically grounded integration in which AI supports rather than supplants professional judgement. Low trust, by contrast, increases the likelihood of resistance, superficial compliance, or over-reliance on AI without critical engagement (Bryk & Schneider, 2002; Fullan et al., 2024).

By delineating these processes, the framework explains how trust shapes the translation of AI into organisational outcomes and accounts for divergent implementation trajectories across schools. It advances existing research by conceptualising trust not as a background condition but as an active mediating mechanism.

## Propositions

To enable empirical examination, the following propositions are advanced:

P1: Trust in leadership positively influences educators' acceptance of AI in school contexts.

P2: Trust in leadership and relational trust positively shape the interpretation of AI outputs as credible and professionally meaningful.

P3: Trust mediates the relationship between AI implementation and organisational outcomes through acceptance, interpretation, and enactment.

P4: Low levels of trust increase the likelihood of resistance, superficial compliance, or uncritical reliance on AI systems.

P5: Integrated leadership practices strengthen trust and enhance the quality of AI enactment.

By specifying trust as a mediating mechanism operating through acceptance, interpretation, and enactment, the framework advances beyond prior models that treat trust as a general facilitator of technology adoption or as a static organisational condition. Instead, it provides a process-based explanation that captures how AI outcomes are actively constructed through relational dynamics within schools (see **Figure 1**).

# DISCUSSION

This discussion interprets the proposed framework in relation to existing theory and examines its implications for educational leadership and research. The central argument is that AI integration in school leadership should not be understood as a purely technical or administrative development, but as a relationally mediated organisational process. While existing scholarship demonstrates that AI can enhance administrative efficiency, support data-informed decision-making, and strengthen strategic planning, it also acknowledges important ethical and governance challenges (Adams & Thompson, 2025; Dai et al., 2025; Fu & Weng, 2024; García-López & Trujillo-Liñán, 2025; Williamson et al., 2024). However, these perspectives have largely remained fragmented, offering limited insight into how such opportunities and challenges are translated into practice within school contexts.

What this study adds is a process-based explanation of how these dynamics are mediated through trust. A key contribution lies in reconceptualising trust as an active mediating mechanism rather than a contextual or enabling condition. Existing research in educational leadership has consistently shown that trust supports

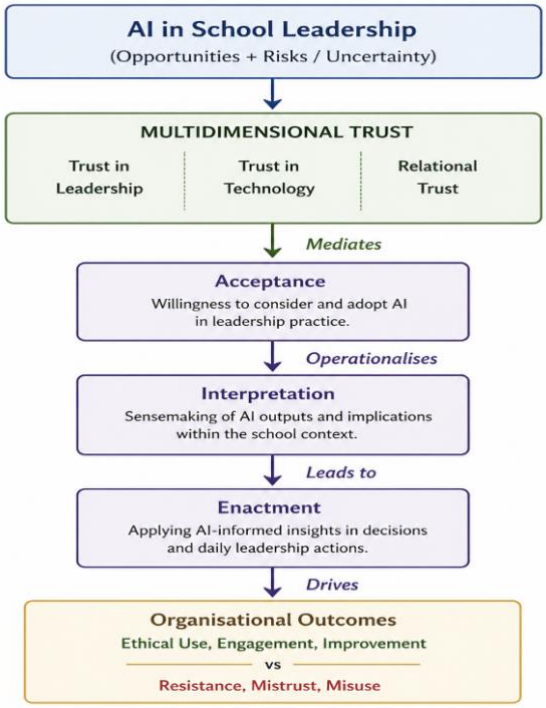


**Figure 1.** Trust as a mediating mechanism in AI-enabled school leadership (Source: Created by the authors)

collaboration, professional learning, openness to change, and school improvement (Bryk & Schneider, 2002; Tschannen-Moran & Gareis, 2015). However, such work has typically treated trust as a background condition, without specifying the mechanisms through which it shapes the implementation of innovation. In contrast, this study demonstrates that, in AI-enabled contexts, trust actively structures how educators engage with technological systems through the processes of acceptance, interpretation, and enactment.

The framework further advances theoretical clarity by specifying how trust operates across these processes. By distinguishing acceptance (initial willingness to engage), interpretation (sensemaking of outputs), and enactment (practical use), the study moves beyond general assertions that "trust matters" and offers a more precise account of how trust shapes organisational processes. This distinction is theoretically significant, as it highlights that challenges in AI implementation may emerge at different stages, for example, resistance at the point of acceptance or misalignment during interpretation. In doing so, the framework aligns with broader organisational research emphasising the importance of sensemaking in shaping responses to complex and uncertain innovations (Weick, 1995).

Another important contribution lies in positioning leadership as a form of relational governance in AI-enabled environments. Current debates on AI in education have largely focused on technical safeguards, regulatory frameworks, and ethical guidelines. While these approaches are necessary, they are insufficient in isolation. AI implementation in school contexts is fundamentally shaped by relationships, professional norms, and interpretive practices. Leadership, therefore, plays a critical role not only in implementing AI but in shaping the relational conditions under which it is understood and used. This perspective aligns with research emphasising that leadership effectiveness is context-dependent and socially embedded (Hallinger, 2018), as well as work highlighting the central role of trust, care, and organisational learning in educational settings (Louis & Murphy, 2017).

The framework also challenges deterministic assumptions about technological innovation in education. Much of the discourse surrounding AI is implicitly underpinned by technological optimism, assuming that enhanced analytical capacity will lead to improved decision-making and organisational outcomes. This study problematises such assumptions by demonstrating that technological potential is contingent upon relational conditions. In high-trust environments, AI may support professional engagement, ethical decision-making, and organisational learning; in low-trust contexts, it may amplify scepticism, resistance, or superficial compliance. This insight contributes to a more nuanced understanding of educational change in Education 4.0, emphasising that social infrastructure remains as critical as technological capability (Edmondson, 2018).

Furthermore, the study extends emerging research linking trust to organisational innovation and learning. Recent work suggests that trust is closely associated with collective teacher innovativeness and the capacity of schools to engage with change (Dedering & Pietsch, 2025). The present framework builds on this perspective by demonstrating that, in AI-enabled environments, trust not only supports innovation but shapes the quality and direction of its enactment. In this sense, trust functions as both an ethical and strategic resource, mediating how technological change is experienced and implemented in practice.

## Limitations and Boundary Conditions

As a conceptual model, the framework does not provide empirical validation and may not fully capture the complexity of AI implementation in diverse contexts. It is primarily applicable to school-level (K–12) settings characterised by moderate AI maturity and assumes relational dynamics typical of many public education systems. Boundary conditions include variations in leadership autonomy, cultural differences in trust formation, and stages of AI adoption. In highly centralised systems or early-stage AI environments, additional factors such as policy constraints, limited infrastructure, or external accountability pressures may influence the proposed relationships.

Alternative perspectives could usefully complement this relational view. Critical and socio-material approaches, for example, could examine power dynamics, material agency in AI–human assemblages, and how broader policy or infrastructural conditions shape (or constrain) trust processes. These perspectives would extend the framework by addressing issues of equity, surveillance, and the political economy of AI in education that lie beyond its current focus.

The framework opens important avenues for empirical research. Future studies could employ structural equation modelling to test the proposed mediation pathways, with acceptance, interpretation, and enactment examined as sequential or parallel processes. Mixed-methods designs and longitudinal case studies would be particularly valuable in capturing how trust evolves as AI becomes embedded in school practice. Comparative research across governance systems and levels of AI maturity would further clarify contextual influences on the model.

To enhance empirical applicability, future research should operationalise the framework's core constructs using established and context-sensitive measures. Trust in leadership may be assessed through validated scales capturing perceptions of competence, integrity, and benevolence (Tschannen-Moran & Gareis, 2015), while trust in technology may be operationalised through perceived reliability, transparency, and fairness (Shin, 2021). Relational trust may be measured through indicators of collaboration, mutual respect, and professional community. The mediating processes can be captured through measures of acceptance (willingness to engage), interpretation (sensemaking and perceived credibility), and enactment (extent and manner of use). Such operationalisation would enable both quantitative testing and qualitative exploration of process dynamics.

Overall, this study reframes AI implementation in school leadership as a relational and organisational process rather than a purely technical or regulatory challenge. By positioning trust as a mediating mechanism, it shifts

the field from viewing AI adoption as a matter of capability or compliance toward understanding it as a process of organisational enactment shaped by leadership and relational conditions. This perspective offers a more theoretically grounded and practically relevant account of how AI becomes meaningful in educational contexts.

# IMPLICATIONS

## Implications for School Leadership Practice

This study has important implications for school leaders navigating the integration of artificial intelligence (AI) within increasingly complex and ethically sensitive educational environments. The central implication is that AI implementation is not primarily a technical challenge but a relational one. While existing guidance often emphasises infrastructure, training, and compliance, the present framework suggests that the success of AI depends fundamentally on leaders' ability to cultivate trust as a precondition for its acceptance and responsible use.

In practical terms, this requires leaders to prioritise transparency, fairness, and professional dialogue when introducing AI systems. Transparency involves communicating not only the purposes of AI but also its limitations, uncertainties, and potential risks. Research has consistently shown that openness and clarity in leadership communication are central to building trust and fostering engagement in change processes. In AI contexts, this becomes even more critical, as educators must be able to understand how data is used and how decisions are generated in order to engage meaningfully with the technology.

Fairness is equally important, particularly given concerns about algorithmic bias and inequitable outcomes. Leaders must ensure that AI systems are implemented in ways that are perceived as just and aligned with professional values. This includes creating opportunities for teachers to question, challenge, and interpret AI-generated insights, thereby reinforcing professional agency rather than undermining it. Such practices align with broader arguments that ethical AI governance in education requires not only technical safeguards but also participatory and context-sensitive leadership.

Moreover, the findings highlight the importance of relational engagement in AI implementation. Leaders should actively involve teachers in the adoption and adaptation of AI tools, fostering a sense of shared ownership and collective responsibility. By embedding AI within collaborative structures, leaders can strengthen trust and ensure that technological change is aligned with pedagogical priorities and professional expertise.

## Implications for Policy and AI Governance

At the policy level, the study suggests that current approaches to AI governance in education may be insufficient if they focus predominantly on regulation, compliance, and technical standards. While such mechanisms are necessary, they do not fully address the relational dynamics that shape how AI is enacted in practice. The framework proposed here indicates that trust should be treated as a central dimension of AI governance, alongside transparency, accountability, and ethical oversight.

This has several implications. First, policy frameworks should explicitly recognise the role of school leadership in mediating AI implementation. Rather than assuming that policy directives translate directly into practice, policymakers should account for the interpretive work carried out by school leaders and the relational conditions that influence this process. Second, policies should encourage participatory approaches to AI adoption, ensuring that teachers and school communities are actively involved in decision-making processes. This can help build trust and enhance the legitimacy of AI initiatives.

Third, there is a need to move toward human-centred AI governance models that integrate technical, ethical, and relational considerations. By embedding trust within governance frameworks, policymakers can create conditions that support both innovation and ethical responsibility.

## Implications for Research

The conceptual framework proposed in this study opens several avenues for future research. First, empirical studies are needed to test the mediating role of trust in AI implementation across different educational contexts. Such research could examine whether the processes of acceptance, interpretation, and enactment function as distinct or sequential pathways, and how these processes interact over time. Structural equation modelling or longitudinal designs would be particularly suitable for testing the proposed mediation pathways and propositions.

Second, future research could explore the relative importance of the three dimensions of trust (trust in leadership, trust in technology, and relational trust) and their interactions in shaping AI adoption. Understanding these dynamics would provide a more nuanced account of AI implementation and contribute to the development of more effective leadership strategies.

Third, comparative studies across centralised and decentralised education systems could shed light on how contextual factors influence the role of trust in AI-enabled leadership. Longitudinal and mixed-method studies would also allow researchers to examine how trust evolves dynamically as AI becomes embedded in school practice, moving beyond static conceptualisations.

# CONCLUSION

This paper has argued that the integration of artificial intelligence (AI) into school leadership cannot be fully understood through a purely technological or regulatory lens. Instead, it must be conceptualised as a relationally mediated process, in which trust plays a central role in shaping how AI is accepted, interpreted, and enacted within educational organisations.

By positioning trust as a mediating mechanism, the study makes a key contribution to the literature on Education 4.0 and educational leadership. It demonstrates that the impact of AI is not determined solely by its technical capabilities but by the relational conditions under which it is implemented. In this sense, trust functions as a form of relational governance, enabling schools to navigate the opportunities and ethical tensions associated with AI in ways that support professional engagement, ethical decision-making, and sustainable improvement.

More broadly, the study challenges deterministic assumptions about technological innovation in education. It suggests that the success of AI depends not only on the sophistication of the technology but on the capacity of leaders to build trust, foster collaboration, and support professional judgement. Without trust, AI risks becoming a source of resistance, mistrust, and superficial compliance; with trust, it has the potential to enhance leadership practice and contribute meaningfully to school improvement.

In conclusion, the paper calls for a shift in how AI in school leadership is conceptualised and studied. Rather than focusing solely on what AI can do, future research and practice should attend to the relational conditions that determine what AI becomes in practice. By positioning trust at the centre of AI-enabled school leadership, this study offers a theoretically grounded framework that reframes technological innovation as a relational achievement, thereby advancing both the scholarship and practice of educational leadership in the context of Education 4.0.

**Author contributions:** NK: Conceptualization, methodology, formal analysis, investigation, writing - original draft, project administration. ME: Conceptualization, resources, writing - review & editing, supervision.

**Funding:** No funding source is reported for this study.

**Declaration of interest**: There was no conflict of interest by the authors.

**Ethical statement:** Not applicable. This is a conceptual article based on a literature review; no human participants were involved.

**AI statement:** The authors used the AI tool Gemini (Google) for translation, improving the language fluency, and editing the manuscript. After using this AI tool, we reviewed and verified the final version of our work. We, as the authors, take full responsibility for the content of our published work.

**Data sharing statement:** Not applicable. No new data were created or analyzed in this study.

# REFERENCES

Adams, D., & Thompson, P. (2025). Transforming school leadership with artificial intelligence: Applications, implications, and future directions. *Leadership and Policy in Schools*, *24*(1), 77–89. https://doi.org/10.1080/15700763.2024.2411295

Arar, K., Tlili, A., Schunka, L., Salha, S., & Saiti, A. (2025). Reimagining educational leadership and management through artificial intelligence: An integrative systematic review. *Leadership and Policy in Schools*, *24*(1), 4–26. https://doi.org/10.1080/15700763.2025.2451982

Bligh, M.C. (2017). Leadership and Trust. In: J. Marques & S. Dhiman (Eds.*), Leadership today (pp. 21-42)*. *Springer Texts in Business and Economics*. Springer, Cham. https://doi.org/10.1007/978-3-319-31036-7_2

Bryk, A. S., & Schneider, B. (2002). *Trust in schools: A core resource for improvement*. Russell Sage Foundation. https://www.jstor.org/stable/10.7758/9781610440967

Bukko, D., Liu, K., & Johnson, A. H. (2021). Principal practices that build and sustain trust: Recommendations from teachers in a high-trust school. *Planning and Changing*, *50*(1/2), 58-74. https://eric.ed.gov/?id=EJ1296864

Caldwell, C., Hayes, L. A., Bernal, P., & Karri, R. (2008). Ethical stewardship–implications for leadership and trust. *Journal of Business Ethics*, *78*(1), 153-164. https://doi.org/10.1007/s10551-006-9320-1

Choung, H., David, P., & Ross, A. (2023). Trust in AI and its role in the acceptance of AI technologies. *International Journal of Human–Computer Interaction*, *39*(9), 1727–1739. https://doi.org/10.1080/10447318.2022.2050543

Connelly, B. L., Crook, T. R., Combs, J. G., Ketchen, D. J., & Aguinis, H. (2018). Competence- and Integrity-Based Trust in Interorganizational Relationships: Which Matters More? *Journal of Management*, *44*(3), 919-945. https://doi.org/10.1177/0149206315596813

Cranston, J. (2011). Relational trust: The glue that binds a professional learning community. *Alberta Journal of Educational Research*, *57*(1), 59-72. https://doi.org/10.55016/ojs/ajer.v57i1.55455

Dai, R., Thomas, M. K. E., & Rawolle, S. (2025). The roles of AI and educational leaders in AI-assisted administrative decision-making: A proposed framework for symbiotic collaboration. *The Australian Educational Researcher*, *52*(2), 1471–1487. https://doi.org/10.1007/s13384-024-00771-8

Dedering, K., & Pietsch, M. (2025). School leader trust and collective teacher innovativeness: On individual and organisational ambidexterity's mediating role. *Educational Review*, *77*(2), 351–380. https://doi.org/10.1080/00131911.2023.2195593

Edmondson, A. C. (2018). *The fearless organization: Creating psychological safety in the workplace for learning, innovation, and growth*. Wiley. https://www.wiley.com/en-us/The+Fearless+Organization%3A+Creating+Psychological+Safety+in+the+Workplace+for+Learning%2C+Innovation%2C+and+Growth-p-9781119477266

Edwards-Groves, C., Grootenboer, P., & Ronnerman, K. (2020). Facilitating a culture of relational trust in school-based action research: Recognising the role of middle leaders. In C. Edwards-Groves, P. Grootenboer, & K. Ronnerman (Eds.), *Partnership and recognition in action research* (pp. 53–70). Routledge. https://doi.org/10.4324/9780429444593

Ellonen, R., Blomqvist, K., & Puumalainen, K. (2008). The role of trust in organisational innovativeness. *European Journal of Innovation Management*, *11*(2), 160-181. https://doi.org/10.1108/14601060810869848

Fu, Y., & Weng, Z. (2024). Navigating the ethical terrain of AI in education: A systematic review on framing responsible human-centered AI practices. *Computers and Education: Artificial Intelligence*, *7*, Article 100306. https://doi.org/10.1016/j.caeai.2024.100306

Fullan, M., Azorín, C., Harris, A., & Jones, M. (2024). Artificial intelligence and school leadership: Challenges, opportunities and implications. *School Leadership & Management, 44*(4), 339–346. https://doi.org/10.1080/13632434.2023.2246856

García-López, I. M., & Trujillo-Liñán, L. (2025). Ethical and regulatory challenges of generative AI in education: A systematic review. *Frontiers in Education*, *10*, Article 1565938. https://doi.org/10.3389/feduc.2025.1565938

Griffiths, A.-J., Alsip, J., Hart, S. R., Round, R. L., & Brady, J. (2021). Together we can do so much: A systematic review and conceptual framework of collaboration in schools. *Canadian Journal of School Psychology*, *36*(1), 59-85. https://doi.org/10.1177/0829573520915368

Hallinger, P. (2018). Bringing context out of the shadows of leadership. Edu*cational Management Administration & Leadership, 46*(1), 5–24. https://doi.org/10.1177/1741143216670652

Handford V., & Leithwood K. (2013). Why teachers trust school leaders. *Journal of Educational Administration*, *51*(2), 194–212. https://doi.org/10.1108/09578231311304706

Keravnos, N., & Symeou, L. (2026). Practices, behaviours and actions of highly trusted primary school principals: An in-depth examination. *Educational Management Administration & Leadership*, *54*(2), 464–495. https://doi.org/10.1177/17411432241262388

Keravnos, N., Lipsou, E., & Pavlakis, M. (2025). Building high faculty trust through leadership integration in primary schools: The role of transformational, instructional and distributed leadership. *Educational Management Administration & Leadership*. https://doi.org/10.1177/17411432251398349

Kesim, E., Atmaca, T., & Turan, S. (2025). Reshaping school cultures: AI's influence on organizational dynamics and leadership behaviors. *Leadership and Policy in Schools*, *24*(1), 117–136. https://doi.org/10.1080/15700763.2024.2413703

Kutsyuruba, B., & Walker, K. (2015). The lifecycle of trust in educational leadership: an ecological perspective. *International Journal of Leadership in Education*, *18*(1), 106–121. https://doi.org/10.1080/13603124.2014.915061

Leithwood, K., Louis, K. S., Anderson, S., & Wahlstrom, K. (2004). *How leadership influences student learning: Review of research*. The Wallace Foundation. https://wallacefoundation.org/sites/default/files/2023-07/How-Leadership-Influences-Student-Learning.pdf

Lipsou E, Keravnos N, Eteokleous N (2026). Artificial intelligence in school leadership: a structured literature review of organisational benefits and ethical challenges. *Artificial Intelligence in Education*, *2*(3), pp. 85–100, doi: https://doi.org/10.1108/AIIE-08-2025-0240

Louis, K. S. (2007). Trust and improvement in schools. *Journal of Educational Change*, *8*(1), 1-24. https://doi.org/10.1007/s10833-006-9015-5

Louis, K. S., & Murphy, J. (2017). Trust, caring and organizational learning: The leader's role. *Journal of Educational Administration*, *55*(1), 103–126. https://doi.org/10.1108/JEA-07-2016-0077

Marks, H. M., & Printy, S. M. (2003). Principal leadership and school performance: An integration of transformational and instructional leadership. *Educational Administration Quarterly*, *39*(3), 370–397. https://doi.org/10.1177/0013161X03253412

Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An integrative model of organizational trust. *The Academy of Management Review*, *20(*3), 709–734. https://doi.org/10.5465/amr.1995.9508080335

Sain, Z. H., Sain, S. H., & Serban, R. (2024). Implementing artificial intelligence in educational management systems: A comprehensive study of opportunities and challenges. *Asian Journal of Managerial Science, 13*(1), 23–31. https://doi.org/10.70112/ajms-2024.13.1.4235

Shin, D. (2021). The effects of explainability and causability on perception, trust, and acceptance: Implications for explainable AI. *International Journal of Human-Computer Studies*, *146*, Article 102551. https://doi.org/10.1016/j.ijhcs.2020.102551

Tschannen-Moran, M. (2014). *Trust matters: Leadership for successful schools* (2nd ed.). Jossey-Bass. https://www.wiley.com/en-us/shop/general-introductory-education/trust-matters-leadership-for-successful-schools-2nd-edition-p-9781118837955

Tschannen-Moran, M., & Gareis, C. R. (2015). Faculty trust in the principal: An essential ingredient in high-performing schools. *Journal of Educational Administration*, *53*(1), 66–92. https://doi.org/10.1108/JEA-02-2014-0024

Wang, Y. (2024). Algorithmic decisions in education governance: Implications and challenges. *Discover Education*, *3*, Article 229. https://doi.org/10.1007/s44217-024-00337-x

Weick, K. E. (1995). *Sensemaking in organizations*. Sage Publications. https://www.sagepub.com/shop/buy-a-book/sensemaking-in-organizations-1-4988

Williamson, B., Molnar, A., & Boninger, F. (2024). *Time for a pause: Without effective public oversight, AI in schools will do more harm than good*. National Education Policy Center. https://nepc.colorado.edu/publication/ai

Yılmaz, O. I. S. (2025). AI, ethical stress, and emotional labor in educational leadership: Toward a human-centered framework. In Ş. Deniz (Ed.), *The new DNA of education: Innovation, technology, equity, and the cognitive turn (pp. 77-131)*. https://doi.org/10.58830/ozgur.pub1137.c4679