WHEN AI CAPABILITY OUTPACES GOVERNANCE: GOVERNANCE–CAPABILITY MISALIGNMENT IN EDUCATIONAL LEADERSHIP

Hasan Mohamed Husain Alaali
DFAS Research Center, Kingdom of Bahrain
**founder@dfas.ai*

Article History
Received: 13/08/2026
Accepted: 09/09/2026
Published online: 21/10/2026

ABSTRACT

Introduction: Artificial intelligence (AI) is expanding from educational support into institutional processes involving administrative, strategic, and consequential decision-making. This development creates a leadership challenge when AI increasingly shapes decision formation while formal authority and accountability remain assigned to human educational leaders (Chen & Ma, 2026; Dai et al., 2025). Although existing literature addresses educational AI governance, leadership, human oversight, decision authority, and adaptive governance, these streams remain insufficiently integrated in explaining changing governance fit (Cavalcante Siebert et al., 2023; Crompton et al., 2026; Jiang & Abdullah, 2026).

Methods: This study employs a structured integrative conceptual review and theory-extension approach to synthesise intersecting literature and extend governance–capability misalignment reasoning to educational leadership.

Results: The theoretical synthesis indicates that governance–capability misalignment can emerge when material changes in AI capability and actual AI decision influence are not accompanied by corresponding adjustments in educational governance configuration. The analysis identifies implications for leadership authority, meaningful human control, intervention capacity, and institutional accountability. Governance recalibration capacity is specified as the mechanism through which institutions can detect emerging divergence and adjust governance arrangements.

Discussion: The study conceptualises educational AI governance as a dynamic alignment problem requiring continuing correspondence between evolving AI capability, substantive decision influence, and institutional governance. It thereby extends educational leadership analysis from the existence of AI governance toward its continuing substantive adequacy.

***Keywords**: Artificial intelligence governance, educational leadership, educational institutions, decision authority, human oversight, accountability, governance capability, governance recalibration.*

INTRODUCTION

AI as Institutional Decision Infrastructure

Artificial intelligence (AI) is moving beyond its earlier role as a primarily instructional or productivity-support technology and is becoming increasingly embedded in the institutional processes through which

educational organisations coordinate, evaluate, plan, and make decisions. Recent research on higher education governance identifies AI applications across strategic planning, administrative decision-making, resource allocation, quality assurance, risk management, institutional data management, and related governance functions (Jiang & Abdullah, 2026). Similarly, the educational leadership literature identifies AI applications spanning administrative efficiency, decision-making and policy formulation, student support, organisational leadership, strategic planning, governance, and compliance (Sposato, 2025). AI governance in education must therefore be considered not only in relation to teaching and learning but also in relation to the institutional infrastructure through which consequential decisions are formed.

This expansion is evident across admissions and applicant screening, student-risk identification, assessment, administrative decision-making, resource allocation, institutional planning, and performance-related processes. Institutional policies have correspondingly expanded from narrow concerns about classroom use toward governance, operational responsibility, privacy, security, accountability, and organisational implementation (An et al., 2025; Chan, 2023). Jiang and Abdullah (2026) characterise AI-enabled governance as an informational infrastructure which reshapes institutional information flows, coordination, accountability, and decision routines.

The resulting transition can be represented as:

AI as Tool → AI as Decision Support → AI as Substantive Decision Influence. This transition does not imply that AI necessarily acquires formal decision-making authority. Rather, it highlights the possibility that AI-generated predictions, classifications, rankings, recommendations, and other outputs may become embedded within recurring institutional processes and thereby exercise increasing influence over how decisions are formed. The central governance issue is consequently no longer simply whether educational institutions adopt AI, but whether AI increasingly becomes embedded in processes through which consequential institutional decisions are shaped.

The Educational Leadership Problem

The increasing institutional role of AI creates a distinctive problem for educational leadership. Educational leaders may retain formal authority and institutional accountability while the informational and analytical processes preceding their decisions become progressively influenced by AI. AI may shape information selection, prediction, recommendation, prioritisation, and ultimately decision formation. Research on AI-assisted administrative decision-making has already demonstrated the importance of distinguishing the complementary roles performed by AI systems and educational leaders rather than treating AI use as straightforward substitution for human decision-makers (Dai et al., 2025). Broader educational leadership research similarly indicates that AI adoption affects leadership practice, organisational decision-making, strategic responsibilities, and the competencies required of educational leaders (Chen & Ma, 2026; Sposato, 2025).

This creates a fundamental distinction:

Formal Decision Authority ≠ Substantive Decision Influence. Formal decision authority identifies the actor institutionally authorised and responsible for making or approving a decision. Substantive decision influence concerns the degree to which an actor, system, or informational process meaningfully shapes the formation and outcome of that decision. An educational leader may therefore retain formal authority while AI-generated outputs exercise increasing substantive influence over the alternatives considered, risks identified, priorities established, or recommendations ultimately adopted.

The leadership problem is consequently not reducible to whether a human remains nominally “in the loop.” It concerns whether formally retained authority remains substantively exercisable as AI-mediated

decision processes evolve. This issue becomes particularly important where leadership responsibility remains stable while technological capability, institutional reliance, or the practical influence of AI outputs changes over time (Chen & Ma, 2026; Dai et al., 2025; Kafa, 2025; Marrone et al., 2025).

Existing Knowledge

The literature already provides substantial foundations for understanding responsible AI governance in education. Responsible AI research has established principles concerning fairness, transparency, accountability, privacy, safety, and human-centred governance. Educational research has translated many of these concerns into institutional AI policies and governance arrangements. Chan (2023), for example, integrates governance and operational dimensions into an AI policy framework for higher education, while more recent work has examined institutional policies, implementation arrangements, and stakeholder responsibilities.

Human oversight and accountability are also well established rather than absent from the literature. Crompton et al. (2026), drawing on expert perspectives across 22 countries and locations, identify human oversight and accountability among the core areas required for higher-education generative-AI governance. Research on meaningful human control further demonstrates that substantive human control requires more than formal human presence and must be considered within the wider socio-technical distribution of responsibility and control (Cavalcante Siebert et al., 2023). Legal scholarship similarly demonstrates that the effectiveness of human oversight depends on questions concerning what is overseen, when oversight occurs, and who possesses the relevant authority (Enqvist, 2023).

Existing scholarship therefore already addresses responsible AI principles, institutional governance, educational AI policies, human oversight, meaningful human control, accountability, human–AI role allocation, educational leadership capabilities, and continuing governance adaptation (Cavalcante Siebert et al., 2023; Chan, 2023; Crompton et al., 2026; Dai et al., 2025; Polat et al., 2025). The theoretical problem addressed in this study does not arise from the absence of these components.

Research Gap

The remaining problem concerns their dynamic integration. Existing educational research increasingly explains how AI should be governed, what responsibilities educational leaders retain, how human and AI roles may be allocated, and why human oversight and accountability matter. Yet these streams provide a less integrated explanation of what occurs when the AI capabilities operating inside educational institutions and their substantive influence over decisions change after governance arrangements have already been established.

Recent evidence makes this issue increasingly important. Jiang and Abdullah (2026) find that AI is becoming embedded in strategic, administrative, and risk-related governance domains while transparency, accountability, trust, equity, workload redistribution, and governance reconfiguration remain comparatively underexamined. Educational leadership scholarship likewise recognises that AI alters leadership environments and requires continuing adaptation (Chen & Ma, 2026), while research on human–AI decision-making and meaningful human control demonstrates that formal human authority alone cannot establish substantive control (Cavalcante Siebert et al., 2023; Dai et al., 2025).
The unresolved relationship can therefore be represented as:

AI Capability Change → Change in Actual AI Decision Influence → Persistence of Existing Governance Configuration → Potential Governance–Capability Misalignment. This relationship builds on AI Governance–Capability Misalignment (AI-GCM), which conceptualises governance effectiveness in terms of continuing correspondence between evolving AI capabilities and the governance configuration surrounding them (Alaali, 2026b). AI-GCM distinguishes changing governance fit from governance that was inadequate from the outset and proposes that initially

substantive governance can become misaligned as the governed AI configuration changes (Alaali, 2026b). The present study extends this reasoning into educational leadership rather than introducing human oversight, authority, accountability, or adaptive governance as new concepts.

Central Research Question

The study addresses the following central research question:

1. RQ: How can changes in AI capability and actual AI decision influence generate governance–capability misalignment in educational institutions, and through what governance mechanism can educational leadership restore alignment?

Two supporting analytical questions refine the inquiry:

1. AQ1. How does actual AI decision influence differ from formally assigned AI decision authority?
2. AQ2. What implications can governance–capability misalignment have for educational leadership authority, substantive human oversight, intervention capacity, and institutional accountability?

Study Objective

The objective of this study is to develop a research-grounded theoretical extension explaining how evolving AI capability and actual decision influence can generate governance–capability misalignment in educational institutions and how educational governance recalibration can contribute to restoring alignment.

Contribution

The central contribution is the extension and domain specification of governance–capability misalignment reasoning within educational leadership. The study conceptualises educational AI governance as a dynamic alignment problem between evolving AI capability and decision influence, on the one hand, and institutional governance configuration, on the other. This moves analysis beyond the existence of governance mechanisms toward the continuing correspondence between what AI can do, how much influence it actually exercises, and the governance arrangements through which educational leaders retain authority and accountability.

Two supporting contributions follow. First, the study explicitly separates formal AI decision role from actual AI decision influence, extending human–AI role-allocation research by examining the substantive influence that AI may exercise even where formal human authority remains unchanged (Cavalcante Siebert et al., 2023; Dai et al., 2025). Second, it specifies governance recalibration capacity within the educational domain as a mechanism through which institutions can recognise changes in AI capability or decision influence and translate them into corresponding governance adjustments. This specification is consistent with educational AI governance research emphasising continuing oversight, accountability, institutional adaptation, and responsive governance arrangements (Crompton et al., 2026).

LITERATURE REVIEW AND THEORETICAL POSITIONING

AI Governance in Educational Institutions

Educational AI governance has developed from broad ethical concerns toward increasingly institutionalised approaches to policy, responsibility, implementation, and control. Early debates around educational AI frequently centred on fairness, privacy, transparency, academic integrity, and responsible use. As AI systems have become more deeply embedded in educational processes, however, governance

questions have increasingly concerned how these principles are translated into institutional policies, responsibilities, procedures, monitoring arrangements, and operational practices.

The progression can be expressed as:

Ethical Principles → Institutional Policies → Governance Mechanisms. Chan's (2023) AI Ecological Education Policy Framework illustrates this development by distinguishing pedagogical, governance, and operational dimensions. Its governance dimension incorporates privacy, security, and accountability, while its operational dimension addresses institutional infrastructure and training. More recent research confirms the movement toward institution-level governance. Crompton et al. (2026) identify eight core areas for higher-education GenAI governance, including ethical and responsible use, privacy and protection, equitable access, AI literacy, integration strategy, human oversight and accountability, and institutional support and infrastructure.

Institutional governance also extends beyond formal policy documents. Jiang and Abdullah's (2026) systematic review shows that AI-enabled governance increasingly operates through strategic, administrative, risk-related, and decision-support processes. Their findings position AI as institutional infrastructure affecting coordination, information flows, accountability, and organisational sense-making. Research examining university AI policies similarly demonstrates growing institutional attention to stakeholder responsibilities, permissible use, implementation, and policy adaptation (An et al., 2025).

These developments indicate that responsible AI principles are increasingly being translated into institutionally embedded mechanisms of accountability, transparency, monitoring, oversight, and operational governance (Chan, 2023; Crompton et al., 2026). Educational AI governance therefore already possesses a substantial normative and institutional foundation. The theoretical challenge is increasingly one of continuing governance adequacy as institutional AI capability and use evolve.

AI and Educational Leadership

The growing institutional role of AI places educational leadership at the intersection between technological capability and organisational authority. AI is increasingly relevant to administrative efficiency, decision-making, policy formulation, student support, strategic planning, governance, compliance, and organisational leadership (Sposato, 2025). Consequently, educational leaders must address not only whether AI improves efficiency or analytical capability, but also how its use affects institutional responsibility, judgement, governance, and decision processes.

Recent evidence reinforces this leadership dimension. Research on digital leadership has already identified the challenges educational leaders face in adapting leadership practices to technological change and strengthening the capabilities required for technology-enabled educational environments (Johari et al., 2023). More specifically, Chen and Ma (2026) conclude that AI can enhance educational leadership while simultaneously introducing risks and new leadership demands. Their review highlights the importance of responsible and human-centred approaches rather than technologically deterministic models of leadership. Similarly, studies of school and institutional leadership indicate that successful AI integration depends on leadership capability, organisational readiness, infrastructure, professional knowledge, and the institutional conditions surrounding implementation (Kafa, 2025; Marrone et al., 2025).

Dai et al. (2025) provide a particularly important bridge between AI capability and leadership authority by examining AI-assisted administrative decision-making. Their framework allocates informational and analytical strengths to AI while retaining important value-laden, supervisory, negotiating, and authority-related functions for educational leaders. This reinforces the continuing importance of human leadership

without implying that formal authority determines the actual distribution of influence throughout the decision process.

Educational leadership can therefore be positioned as the institutional interface between:

Technological Capability

And

Decision Authority + Accountability

The central governance question is whether this interface remains substantively aligned as AI becomes more capable, more embedded, or more influential.

Human–AI Decision Authority and Actual Influence

Research on human–AI decision-making has increasingly moved beyond a binary choice between human and machine decision-makers. AI-assisted administrative decision-making, hybrid decision arrangements, decision-support systems, algorithmic recommendations, and human final authority all involve different distributions of analytical input, judgement, responsibility, and control (Dai et al., 2025).

For governance purposes, two concepts require separation.

Formal Decision Authority refers to the institutionally recognised right and responsibility to make, approve, reject, or assume responsibility for a decision.

Actual AI Decision Influence refers to the substantive extent to which AI-generated information, predictions, rankings, recommendations, or other outputs shape the formation or outcome of that decision.

The distinction matters because retaining formal human authority does not establish that human and AI influence remain substantively unchanged. An AI system formally classified as decision support may influence which information receives attention, how alternatives are ranked, which risks are prioritised, or which course of action appears institutionally preferable. Conversely, an institution may deploy technically capable AI while deliberately constraining its actual influence through strong human review and intervention arrangements.

Meaningful human control scholarship supports this distinction by locating responsibility within a socio-technical relationship between human actors and AI systems rather than treating nominal human involvement as sufficient (Cavalcante Siebert et al., 2023). Human-oversight research likewise demonstrates that substantive oversight depends on the institutional design of oversight authority and its practical operation (Enqvist, 2023). The distinction between formal authority and actual influence therefore conceptually bridges existing human–AI decision research and the governance–capability relationship developed later in this study.

Human Control and Dynamic Governance

Human involvement in AI-mediated decision-making varies substantially in its substantive meaning. For analytical purposes, three levels can be distinguished:

Human Participation → Human Oversight → Meaningful Human Control

Human participation exists where a human actor remains procedurally involved in the decision process. Participation alone does not establish that the actor possesses sufficient information, authority, or practical capacity to alter the process.

Human oversight requires a stronger supervisory relationship in which a human actor reviews or monitors AI-mediated activity and is institutionally positioned to identify relevant concerns. Enqvist (2023) demonstrates that oversight must be understood in relation to what is being overseen, when intervention is possible, and who exercises the relevant oversight function.

Meaningful human control requires still more. It concerns whether human actors possess sufficient knowledge, authority, and practical capacity to evaluate, challenge, alter, override, or escalate AI-mediated processes. Cavalcante Siebert et al. (2023) conceptualise meaningful human control as a socio-technical requirement linked to responsibility and actionable control rather than mere human presence. Related work on responsibility gaps demonstrates why assigning responsibility to a human actor becomes problematic where practical control over technologically mediated outcomes is limited (Matthias, 2004).

The temporal dimension intensifies this problem. AI systems, deployment contexts, organisational reliance, and institutional uses may change after governance arrangements are established. Consequently:

Formal Human Participation May Persist
While **Substantive Human Control Changes**
And **Existing Governance Arrangements May Become Less Adequate**

Contemporary educational governance research already recognises the need for continuing institutional adaptation. Crompton et al. (2026) emphasise continuing policy development, oversight, institutional support, and governance processes as GenAI evolves, while institutional policy research similarly identifies the importance of responsive governance arrangements (An et al., 2025). The unresolved issue is therefore not whether governance should adapt. It is how changing AI capability and actual decision influence can generate a specific deterioration in the fit between AI and the governance arrangements surrounding it.

Literature Synthesis and Theoretical Gap

The reviewed streams establish substantial but differently positioned components of educational AI governance. Responsible AI identifies normative expectations; educational AI governance translates these expectations into institutional policies and controls; educational leadership research identifies leadership responsibilities and capabilities; human–AI decision research addresses role allocation; meaningful human control establishes requirements for substantive human intervention; and adaptive governance recognises the need for continuing adjustment (An et al., 2025; Cavalcante Siebert et al., 2023; Chan, 2023; Chen & Ma, 2026; Crompton et al., 2026; Dai et al., 2025; Jiang & Abdullah, 2026).

Table 1
Existing Literature Streams and the Remaining Explanatory Problem

Literature Stream	Established Contribution	Remaining Explanatory Problem
Responsible AI	Ethical principles	Continuing governance fit

Educational AI governance	Policies and institutional controls	Governance response to capability evolution
Educational leadership	Leadership roles and capabilities	Substantive authority under changing AI influence
Human–AI decision-making	Allocation of decision roles	Formal role versus actual influence
Meaningful human control	Conditions for substantive control	Persistence of those conditions as AI changes
Adaptive governance	Need for continuing adjustment	Specific misalignment–recalibration mechanism

The gap lies in the relationship across these streams rather than in the absence of their individual components. The specific theoretical problem concerns whether governance arrangements that were appropriate under one AI capability and influence configuration remain substantively appropriate after that configuration materially changes. Existing research establishes the importance of responsible governance, leadership authority, human oversight, accountability, role allocation, and institutional adaptation. Yet, it provides a less integrated explanation of how changes in AI capability and actual decision influence can alter the continuing fit between these elements.

AI-GCM provides the parent theoretical reasoning for this problem by treating governance effectiveness as dependent on continuing correspondence between AI Capability Configuration and Governance Configuration and by explaining how initially substantive governance can lose fit as AI evolves (Alaali, 2026b). The educational extension specifies this relationship through actual AI decision influence, educational leadership authority, meaningful human control, institutional accountability, and governance recalibration.

The resulting theoretical sequence is:

Capability Evolution
→ Influence Evolution
→ Governance Persistence
→ Governance–Capability Misalignment
→ Governance Recalibration

This sequence defines the unresolved relationship examined in the study and provides the basis for extending AI-GCM into educational leadership and AI-mediated institutional decision-making.

METHODOLOGY

Research Design

This study adopts a **Structured Integrative Conceptual Review and Theory Extension** design. The methodological objective is theoretical integration across intersecting bodies of research rather than estimation of a narrowly defined empirical effect. This approach is appropriate because the focal problem spans several literatures—including educational AI governance, educational leadership, human–AI decision-making, meaningful human control, accountability, and adaptive governance—whose individual contributions are established but whose relationships require further theoretical integration.

Consistent with integrative review methodology, the study does not claim to exhaustively identify every publication in the field. Instead, it systematically identifies theoretically relevant evidence across intersecting literature streams and integrates that evidence to examine conceptual relationships and develop the proposed theoretical extension. The methodology therefore combines structured evidence identification with conceptual synthesis without presenting the study as a systematic literature review.

The theory-extension component builds specifically on AI Governance–Capability Misalignment (AI-GCM), which conceptualises governance adequacy in terms of continuing correspondence between evolving AI capabilities and the governance configuration surrounding them (Alaali, 2026b). The purpose is not to reproduce the parent theory but to examine how its central alignment logic operates within educational leadership and AI-mediated institutional decision-making.

Literature Identification and Search Strategy

Literature identification was conducted through a structured, theory-oriented search across Scopus, Web of Science, ERIC, and Google Scholar. The primary publication period was 2019–2026, reflecting the rapid development of AI governance, generative AI, educational leadership, and AI-mediated institutional decision-making. Earlier foundational studies were retained where they were theoretically necessary for concepts such as responsibility, accountability, human control, and technologically mediated decision-making.

The search strategy was organised around intersecting search families rather than a single narrowly specified query. Search combinations included: “AI governance” AND education; “artificial intelligence” AND “educational leadership”; AI AND “decision authority”; AI AND “human oversight”; AI AND “meaningful human control”; AI AND accountability AND education; AI AND “institutional governance”; AI AND “adaptive governance”; and AI AND “administrative decision-making” AND education.

Because the study employs a structured integrative conceptual review and theory-extension methodology rather than a systematic literature review, literature identification was directed by theoretical relevance and cross-stream conceptual coverage rather than exhaustive enumeration of every publication returned by the databases. The review therefore did not employ a PRISMA-style record-counting procedure based on a fixed corpus of retrieved records, duplicate removal, and sequential title/abstract/full-text exclusion. Accordingly, numerical values for records retrieved, duplicates removed, records screened, records excluded, and final retained records are not reported, because such counts were not generated through the review procedure and would constitute retrospective reconstruction.

Instead, source selection proceeded iteratively through relevance assessment, conceptual coding, and backward and forward citation searching. Sources were retained where they contributed substantively to one or more of the theoretical relationships examined in the study, particularly AI capability, actual decision influence, educational leadership authority, human oversight and intervention, accountability, institutional governance, and governance adaptation. The search and selection process continued until the intersecting literature streams provided sufficient theoretical coverage to evaluate established relationships, identify conceptual tensions, and develop the educational extension of AI Governance–Capability Misalignment theory.

The resulting evidence base therefore represents a theoretically purposive integrative corpus, not a statistically exhaustive sample of the literature. This methodological boundary is consistent with the objective of the study: to integrate established but dispersed theoretical relationships and develop a domain-specific theory extension rather than estimate the prevalence of findings across a systematically enumerated body of studies.

Eligibility Criteria

Studies were eligible for inclusion if they substantively addressed one or more of the theoretical domains relevant to the study: AI governance, educational leadership, institutional AI governance, AI-mediated decision-making, human oversight or meaningful human control, decision authority, accountability, institutional AI capability, or governance adaptation.

Priority was given to studies that contributed directly to understanding relationships among technological capability, institutional decision processes, leadership authority, substantive human control, accountability, and governance adjustment. Studies addressing AI in education without an institutional, leadership, decision-making, or governance dimension were therefore not automatically eligible. Earlier foundational studies outside the primary 2019–2026 publication period were eligible where they provided theoretically indispensable foundations for responsibility, accountability, human control, or AI-mediated decision-making.

Studies were excluded where their primary focus was technical AI performance, student acceptance without governance implications, pedagogical effectiveness without institutional-governance relevance, generic commentary on AI, or topics without substantive relevance to the theoretical relationships under investigation. Duplicate records, where identified during the documented search process, were retained only once.

Source Selection, Quality, and Contextual Coverage

Source selection prioritised peer-reviewed journal research, systematic and integrative reviews, contemporary empirical studies, theoretically relevant conceptual research, and authoritative institutional evidence where required for institutional anchoring. Particular emphasis was placed on literature published between 2023 and 2026, given the rapid evolution of generative AI, institutional AI adoption, educational governance, and human–AI decision arrangements during this period (Chen & Ma, 2026; Crompton et al., 2026; Jiang & Abdullah, 2026).

Contextual coverage was considered during source selection. The review incorporated evidence from higher education and school settings and from diverse geographical and institutional contexts where theoretically relevant, high-quality evidence was available. Contextual diversity was used to strengthen the domain analysis rather than as a substitute for theoretical relevance or source quality.

Conceptual Coding and Analysis

The analysis combined deductive and inductive conceptual coding. Deductive codes were derived from the research question, the established AI-GCM theoretical problem, and concepts identified in prior governance and human-control research (Alaali, 2026b; Cavalcante Siebert et al., 2023). Inductive coding captured recurrent relationships, tensions, and mechanisms emerging from the retained literature.

Six primary coding domains guided the analysis: AI capability; actual decision influence; decision authority; human oversight and intervention; accountability; and governance adaptation.

Coded extracts from retained sources were compared within and across coding domains to identify recurring patterns, conceptual tensions, established relationships, and underexplored relationships relevant to the theoretical extension. Concepts were assigned per theoretical function to reduce construct overlap and double-counting. Particular attention was given to distinguishing AI capability from actual AI decision influence, formal authority from substantive control, and the formal presence of governance mechanisms from continuing governance–capability fit.

A literature-synthesis matrix subsequently linked the evidence through the analytical sequence:

Constructs → Relationships → Mechanisms → Propositions

This structure provided the basis for proposition development without treating recurrent terminology alone as evidence of a theoretically distinct construct.

Theory-Extension Procedure
Theory extension proceeded through a staged analytical sequence:

Literature Identification → Conceptual Coding → Identification of Established Relationships → Cross-Stream Tension Analysis → Identification of Unresolved Relationship → Comparison with AI-GCM Reasoning → Educational Leadership Domain Specification → Construct-Role Separation → Mechanism Development → Proposition Development → Boundary Conditions → Future Empirical Tests

AI-GCM provided the parent theoretical architecture for examining changing governance–capability correspondence (Alaali, 2026b). The reviewed educational literature was then used to specify how this relationship operates within educational leadership and institutional decision processes, particularly through actual AI decision influence, educational leadership authority, meaningful human control, institutional accountability, and governance recalibration. Each construct was assigned a distinct theoretical role to preserve causal and conceptual separation while maintaining continuity with the parent theory.

THEORETICAL DEVELOPMENT

AI Governance–Capability Misalignment as the Parent Theoretical Position
The theoretical starting point of this study is AI Governance–Capability Misalignment (AI-GCM). AI-GCM conceptualises governance adequacy as relational and temporally situated rather than as a fixed property established merely through the presence of governance policies or controls. Governance arrangements may initially correspond appropriately to an AI system but subsequently lose substantive fit as AI capabilities, organisational uses, decision influence, autonomy, or operational reach change (Alaali, 2026b).

Within the parent theory, Governance Configuration encompasses the temporally situated arrangement of authority, accountability, monitoring, constraint, audit, intervention, escalation, and reconfiguration surrounding AI (Alaali, 2026b). This understanding is consistent with research demonstrating that meaningful human control and oversight depend on substantive relationships among responsibility, authority, information, and intervention capacity rather than formal human presence alone (Cavalcante Siebert et al., 2023; Enqvist, 2023; Lazcoz & de Hert, 2023). Institutional AI governance research similarly emphasises human oversight, accountability, monitoring, and continuing policy adaptation (Crompton et al., 2026).

AI-GCM therefore shifts the analytical question from whether AI governance exists to whether governance continues to correspond substantively to the AI configuration it governs. It further distinguishes governance–capability misalignment from pre-existing governance weakness, symbolic governance, implementation failure, or capability deficiency. Governance may have been substantive and initially appropriate yet become progressively inadequate because the governed technological configuration has materially changed (Alaali, 2026b).

The present study extends this theoretical logic into educational leadership and AI-mediated institutional decision-making. Rather than reproducing the complete temporal architecture of AI-GCM, it specifies how changing AI capability and substantive decision influence interact with educational leadership

authority, human oversight, intervention capacity, and institutional accountability. This domain extension is also informed by institutionally embedded governance reasoning emphasising accountability, operational control, traceability, intervention, escalation, and adaptable governance arrangements (Alaali, 2026a).

The core theoretical premise is therefore:

Educational AI governance remains adequate only insofar as the governance configuration continues to correspond substantively to the AI capabilities and actual decision influence operating within institutional decision processes.

AI Capability Configuration (ACC)

AI Capability Configuration (ACC) refers to the configuration of AI capabilities institutionally available and operationally embedded within educational processes at a given point in time. Its theoretical role is that of an antecedent condition.

ACC is retained from the parent AI-GCM architecture but specified for educational institutions (Alaali, 2026b). Relevant capabilities may include generation, prediction, recommendation, prioritisation, evaluation, decision support, and semi-autonomous execution. Educational AI research demonstrates that such capabilities increasingly operate across institutional governance, administrative decision-making, strategic planning, student support, assessment, and related organisational functions (Chan, 2023; Jiang & Abdullah, 2026; Sposato, 2025).

Within educational institutions, these capabilities may be incorporated into admissions, student-risk identification, assessment, administrative decision-making, resource allocation, institutional planning, and other leadership or governance processes. However, the technical availability of a capability does not establish either its formal decision authority or its substantive influence over institutional decisions. Three analytical questions must therefore remain distinct:

What can the AI do?
≠
What is the AI formally authorised to do?
≠
How much does the AI actually influence the decision?

Accordingly:

AI Capability ≠ AI Authority ≠ AI Decision Influence

This separation distinguishes the technological capabilities available within an educational institution from the authority formally assigned to AI and from the substantive influence AI outputs exercise within decision processes. ACC therefore establishes the technological and operational antecedent against which actual decision influence and continuing governance correspondence can be assessed.

Actual AI Decision Influence (ADI)

Actual AI Decision Influence (ADI) is defined as the substantive extent to which AI-generated information, predictions, recommendations, rankings, or outputs shape the formation, selection, prioritisation, or implementation of an institutional decision, irrespective of formally assigned AI decision authority. Its theoretical role is that of a transmission mechanism through which embedded AI capability acquires substantive relevance for educational governance.

ADI specifies the effective decision-influence logic of AI-GCM within the educational domain. It captures how AI outputs may acquire increasing substantive influence over institutional decisions even where formally allocated decision rights remain unchanged. The educational extension makes this relationship analytically explicit by distinguishing the formal assignment of decision authority from the substantive influence exercised through AI outputs.

Thus: **Formal AI Decision Role ≠ Actual AI Decision Influence**

An AI system may remain formally classified as a recommendation or decision-support mechanism while its outputs acquire substantial practical influence over decisions institutionally attributed to educational leaders. Research on AI-assisted educational administration similarly demonstrates that AI systems and educational leaders can perform distinct but interacting roles in decision formation (Dai et al., 2025).

Operational embedding is the extent to which AI outputs are incorporated into recurring institutional workflows and routinely used as inputs to educational decisions. It provides the principal bridge through which available technological capability can acquire substantive decision influence.

The translation from:

ACC → ADI

may be strengthened by repeated reliance on AI outputs, incorporation of AI recommendations into routine workflows, default or prioritised presentation of AI recommendations, perceived reliability, limited availability of viable human alternatives, decision speed and time pressure, increasing organisational dependence on AI-supported processes, and the practical costs of challenging, overriding, or bypassing AI recommendations. Research concerning educational leadership, AI integration, organisational readiness, and human–AI decision-making supports the relevance of such institutional conditions to the practical embedding of AI in educational decision processes (Dai et al., 2025; Kafa, 2025; Marrone et al., 2025).

These conditions are treated as capability-to-influence translation mechanisms rather than additional core constructs. Their analytical purpose is to explain why comparable technological capabilities may generate different levels of substantive AI influence across educational institutions.

P1. Greater operational embedding of AI capability in educational decision processes is associated with greater potential for actual AI decision influence, even where formally assigned AI decision authority remains unchanged.

This proposition is consistent with meaningful human control research demonstrating that formal responsibility alone does not establish substantive distributions of influence and control (Cavalcante Siebert et al., 2023), while specifying within education the AI-GCM distinction between formal governance arrangements and effective AI decision influence (Alaali, 2026b).

Educational Governance Configuration (EGC)

Educational Governance Configuration (EGC) is defined as the temporally situated institutional arrangement through which decision rights, leadership authority, accountability, monitoring, oversight obligations, intervention rights, escalation procedures, traceability, and governance reconfiguration are organised around AI-mediated educational decisions. Its theoretical role is a conditioning governance architecture.

EGC is an educational-domain specification of Governance Configuration within AI-GCM rather than an independent construct disconnected from the parent theory. It contextualises the governance functions of the parent theory around educational leadership structures, institutional responsibilities, and educational decision processes.

This specification is consistent with educational AI governance research emphasising the translation of responsible-AI principles into institutional policies, accountability arrangements, human oversight, implementation processes, and continuing governance mechanisms (An et al., 2025; Chan, 2023; Crompton et al., 2026). These developments reinforce the importance of embedding governance requirements within the organisational arrangements through which AI-mediated educational activities are authorised, monitored, and overseen.

EGC therefore extends beyond the formal existence of an institutional AI policy. It concerns whether educational leaders possess substantive governance mechanisms through which they can:

Authorise → Monitor → Review → Challenge → Intervene → Escalate → Reconfigure

The relevant question is therefore not merely:

Does the educational institution have AI governance?

but: Does its current governance configuration remain appropriate for the capabilities and substantive decision influence of the AI systems it currently uses?

The distinction shifts the analytical focus from governance presence to continuing governance fit.

Educational Governance–Capability Misalignment

Educational Governance–Capability Misalignment is a domain-specific manifestation of AI-GCM in which the governance configuration surrounding AI-mediated educational decisions no longer corresponds substantively to the AI capabilities and actual decision influence operating within those decisions. Its theoretical role is the central theoretical condition connecting changes in technological capability and decision influence to educational governance consequences.

This condition does not constitute a theory separate from AI-GCM. The parent theory identifies Capability, Authority, Visibility, and Control Misalignment as distinct but potentially interacting manifestations of deteriorating correspondence between current AI Capability Configuration and Governance Configuration (Alaali, 2026b). The educational extension specifies how this broader condition can emerge within educational leadership and institutional decision processes.

The focal relationship is:

Material Change in ACC
→
Change in ADI

- **Persistence of Existing EGC**

→
Potential Educational Governance–Capability Misalignment

The relationship is conditional than deterministic. Increasing AI capability does not itself constitute misalignment, and increasing actual AI decision influence does not necessarily imply governance

failure. Misalignment emerges when material changes in capability and influence are not accompanied by governance adjustments sufficient to preserve substantive correspondence between the AI configuration and the arrangements intended to authorise, oversee, constrain, intervene in, and assign accountability for its institutional use.

The educational extension therefore focuses on whether changes in ACC and ADI become sufficiently material that existing arrangements of leadership authority, human oversight, intervention capacity, traceability, and accountability cease to correspond adequately to the AI-mediated decision process. Contemporary educational governance research demonstrates expanding AI involvement in institutional governance alongside continuing requirements for accountability, oversight, and governance adaptation (Crompton et al., 2026; Jiang & Abdullah, 2026). Research on meaningful human control and educational decision-making likewise establishes the importance of maintaining substantive human authority and intervention capacity as technologically mediated decision processes evolve (Cavalcante Siebert et al., 2023; Dai et al., 2025).

P2. Governance–capability misalignment is more likely when material changes in AI capability and actual AI decision influence occur without corresponding adjustment in educational governance configuration.

P2 specifies the parent AI-GCM relationship within educational leadership without reproducing its complete temporal architecture. The fuller parent theory separately addresses Relative Velocity Gap, Accumulated Temporal Divergence, organisational lags, governance reconfiguration, and temporary governance–capability fit (Alaali, 2026b). The narrower educational architecture isolates the relationships required to explain how governance–capability misalignment emerges within AI-mediated educational decision-making and provides the basis for examining its consequences and subsequent governance recalibration.

GOVERNANCE CONSEQUENCES AND EDUCATIONAL RECALIBRATION

Leadership Authority Divergence

Governance–capability misalignment can divert between formally assigned educational leadership authority and the substantive influence exercised within AI-mediated decision processes. This consequence does not require a formal transfer of decision authority from educational leaders to AI systems. Formal authority may remain institutionally unchanged while AI-generated predictions, classifications, recommendations, rankings, or prioritisation increasingly shape how decisions are formed.

Research on AI-assisted educational administration identifies differentiated but interacting roles for AI systems and educational leaders rather than treating decision authority as a simple human–machine binary (Dai et al., 2025). Educational leadership research similarly indicates that AI increasingly affects decision processes, organisational responsibilities, strategic activity, and leadership practice while educational leaders remain institutionally responsible for AI-mediated organisational activity (Chen & Ma, 2026; Marrone et al., 2025).

Within AI-GCM, this condition corresponds to Authority Misalignment, in which effective AI decision influence no longer corresponds to formally allocated decision rights or accountability (Alaali, 2026b). Leadership authority divergence therefore specifies the educational consequence of this parent-theory relationship.

The distinction can be expressed as:

Formally Retained Leadership Authority ≠ Substantively Unchanged Decision Influence

A principal, administrator, academic leader, committee, or other authorised actor may continue to possess the institutional right to approve a decision while the informational architecture preceding that approval increasingly determines which alternatives are visible, how risks are assessed, and which outcomes appear preferable. The governance concern is therefore whether formally retained educational leadership authority remains substantively exercisable within the actual decision process.

P3. Higher governance–capability misalignment is associated with a greater likelihood of divergence between formally assigned educational leadership authority and the substantive influence exercised by AI outputs.

Human Oversight and Meaningful Control

Leadership authority divergence also raises questions concerning the substantive quality of human involvement. Human participation should not be treated as equivalent to effective oversight or meaningful control. Three progressively stronger conditions can be distinguished:

Human Participation → Human Oversight → Meaningful Human Control

Human participation indicates that a human actor remains procedurally involved. Human oversight requires the capacity to review or supervise an AI-mediated process. Meaningful human control requires a stronger combination of relevant information, authority, and practical intervention capacity sufficient to affect the process or its outcome (Cavalcante Siebert et al., 2023; Veluwenkamp, 2022).

For educational leadership, substantive control can be represented through the practical capacity to:

Evaluate → Challenge → Override → Escalate

Legal and governance scholarship reinforces this distinction. Human oversight cannot be assessed merely by identifying a nominal human decision-maker; its effectiveness depends on what is being overseen, when intervention remains possible, and which actor possesses the authority and practical capacity to intervene (Enqvist, 2023). Similarly, lifecycle-oriented human governance requires more than formally inserting human actors into technologically mediated processes (Lazcoz & de Hert, 2023).

Governance–capability misalignment can therefore coexist with continued procedural human participation. Educational leaders may still review or formally approve decisions while increased AI influence, reduced observability, institutional dependence, or inadequate intervention mechanisms diminish their practical capacity to alter how those decisions are formed.

Within AI-GCM, this consequence corresponds principally to **Control Misalignment**, in which safeguards, intervention arrangements, escalation mechanisms, or human-control structures cease to correspond adequately to actual AI capability, autonomy, or operational reach (Alaali, 2026b).

P4. The substantive effectiveness of human oversight depends on educational leaders possessing sufficient authority and practical capacity to evaluate, challenge, override, or escalate AI-mediated decisions.

Accountability–Control Incongruence

A further consequence arises when institutional responsibility remains attached to a human actor whose substantive control over AI-mediated decision formation has weakened.

Accountability–control incongruence is a condition in which formal responsibility for an institutional decision remains assigned to a human actor whose substantive capacity to shape, challenge, or alter the AI-mediated formation of that decision has become comparatively limited.

The relevant distinction is:

Formal Responsibility ≠ Substantive Decision Control

The possibility that responsibility and control can diverge is well established. Responsibility-gap scholarship demonstrates the difficulty of attributing responsibility when autonomous or learning systems weaken conventional relationships between agency and control (Matthias, 2004). Meaningful human control research similarly emphasises that responsibility attributed to human actors should correspond to their substantive capacity to exercise control (Cavalcante Siebert et al., 2023; Veluwenkamp, 2022). AI-GCM incorporates this problem by recognising that increasing AI agency and decision influence can separate formal accountability from substantive control (Alaali, 2026b).

The educational extension specifies how such incongruence can arise through governance–capability misalignment. An educational leader may remain formally accountable for a consequential institutional decision even though the AI-mediated process increasingly shapes the information considered, alternatives prioritised, or recommendation adopted. The governance problem is therefore not accountability allocation alone, but whether assigned accountability continues to correspond to practical decision control.

P5. Greater governance–capability misalignment increases the likelihood of incongruence between institutional accountability and substantive control over AI-mediated decision formation.
Governance Recalibration Capacity (GRC)

Governance–capability misalignment need not persist. Governance Recalibration Capacity (GRC) is defined as an educational institution's capacity to detect material changes in AI capability or actual decision influence and translate that detection into corresponding adjustment of relevant governance arrangements. Its theoretical role is a corrective mechanism.

GRC is an educational-domain specification of the reconfiguration logic established within AI-GCM rather than an independent corrective theory. The parent theory conceptualises governance response through:

Governance Sensing → Governance Seizing → Governance Reconfiguration

Governance sensing identifies deterioration in governance–capability correspondence; governance seizing converts recognition into an authorised organisational response; and governance reconfiguration modifies the relevant governance configuration (Alaali, 2026b).

Within educational institutions, this logic is specified as:

Detect → Assess Governance Relevance → Authorise Response → Reallocate Decision Rights Where Necessary → Adjust Oversight and Intervention → Realign Accountability → Review Restored Fit

This sequence distinguishes awareness from institutional capacity to act. Detection of increased AI influence does not itself restore alignment. Educational institutions must possess the authority, organisational processes, and governance mechanisms necessary to convert detection into substantive

reconfiguration. Contemporary educational AI governance research similarly emphasises continuing policy review, institutional oversight, accountability, and governance adaptation as AI use evolves (An et al., 2025; Crompton et al., 2026; Jiang & Abdullah, 2026).

Recalibration can operate through governance levers within institutional authority, including AI-use permissions, allocation of decision rights, mandatory human review, acceptance or rejection authority, escalation procedures, audit logging, procurement requirements, deployment restrictions, vendor obligations, and temporary suspension of AI use. These mechanisms enable institutions to adjust the governance configuration surrounding AI use without presuming direct control over proprietary model architecture, training data, model weights, or vendor-controlled technical infrastructure.

P6. Greater governance recalibration capacity reduces the likelihood that changes in AI capability and actual AI decision influence develop into persistent governance–capability misalignment.

INTEGRATED MODEL AND INSTITUTIONAL ANCHORING

Integrated Educational Extension of AI-GCM

The preceding theoretical development can be consolidated into an integrated educational extension of AI-GCM. The architecture preserves the distinct theoretical roles of its components rather than treating them as equivalent governance dimensions.

Figure 1
Educational Leadership Extension of AI Governance–Capability Misalignment



This architecture specifies how AI-GCM operates within educational leadership while retaining the parent theory's central premise that governance adequacy depends on continuing correspondence between the AI configuration and the governance configuration surrounding it (Alaali, 2026b).

Interaction Between the Constructs

The five principal constructs perform distinct but interdependent theoretical functions. ACC represents the antecedent technological configuration, while capability-to-influence translation mechanisms explain how available capability can acquire substantive organisational influence through operational embedding, repeated reliance, routinisation, time pressure, and institutional dependence. ADI captures the resulting substantive influence of AI outputs on educational decision formation.

EGC provides the governance architecture against which changes in capability and decision influence are assessed. GCM emerges when this architecture no longer corresponds substantively to the AI capability and actual decision influence operating within institutional decisions. GRC provides the corrective mechanism through which deteriorating correspondence can be detected and relevant governance arrangements reconfigured.

The interaction is therefore conditional than deterministic. Increasing AI capability does not itself generate governance–capability misalignment. Misalignment becomes possible when capability acquires substantive decision influence and the corresponding governance configuration does not adjust sufficiently to preserve substantive fit. This interaction specifies the AI-GCM mechanism within educational leadership while preserving the distinct theoretical functions of ACC, ADI, EGC, GCM, and GRC.

Temporal Logic

The educational extension can be represented through a simplified temporal sequence:

Figure 2

Educational Domain Sequence of Governance–Capability Misalignment and Recalibration

t_0 — Governance–Capability Fit

→ t_1 — Material AI Capability Change

→ t_2 — Change in Actual AI Decision Influence

→ t_3 — Persistence of Existing Educational Governance Configuration

→ t_4 — Potential Governance–Capability Misalignment

→ t_5 — Leadership Authority / Human Control / Accountability Consequences

→ t_6 — Governance-Relevant Detection

→ t_7 — Educational Governance Recalibration

→ t_8 — Potential Restored Governance–Capability Fit

E-ISSN: 2289-9669

This sequence is intentionally narrower than the complete temporal architecture of AI-GCM. The parent theory additionally specifies Relative Velocity Gap, Accumulated Temporal Divergence, Detection Lag, Response Lag, Reconfiguration Lag, Temporary Governance–Capability Fit, Alignment Duration, and organisational learning (Alaali, 2026b). The educational extension retains only those temporal relationships necessary to explain the emergence, consequences, and correction of governance–capability misalignment within educational leadership and institutional decision-making.

The endpoint is potential restored fit rather than permanent equilibrium. AI capability, operational embedding, actual decision influence, and institutional dependence may continue to change after recalibration. Consequently, a reconfigured EGC may restore substantive correspondence temporarily while remaining subject to subsequent reassessment and further governance adjustment.

Boundary Conditions

The institutional consequences of educational GCM are expected to vary according to the characteristics of the decision environment. Misalignment should become more consequential where decision stakes are high, decisions are difficult to reverse, stakeholder exposure is substantial, relevant AI processes are difficult to observe or interpret, or institutional dependence on AI outputs is high.

These conditions do not define GCM and are not additional core constructs. Rather, they condition the significance and potential consequences of an existing misalignment. This reasoning is consistent with research showing that meaningful human oversight becomes particularly important where AI-mediated decisions carry substantial consequences and practical intervention is constrained (Cavalcante Siebert et al., 2023; Enqvist, 2023). Research on AI-assisted educational decision-making similarly indicates that the appropriate allocation of AI and leadership roles depends partly on the nature and implications of the decision involved (Dai et al., 2025).

P7. The institutional consequences associated with governance–capability misalignment are expected to be greater for high-consequence, difficult-to-reverse educational decisions than for low-consequence, readily reversible AI-assisted activities.

This proposition is also consistent with the parent AI-GCM emphasis on operational reach, autonomy, observability, and institutional dependence as conditions affecting the development and consequences of misalignment (Alaali, 2026b).

Documented Institutional Illustrations

The analytical applicability of the educational extension can be illustrated through three documented contexts represented in the reviewed literature.

First, AI-assisted administrative decision-making illustrates the distinction between formal authority and substantive AI influence. Dai et al. (2025) conceptualise AI systems and educational leaders as performing differentiated roles within administrative decision-making. AI can perform informational and analytical functions while educational leaders retain roles involving value-laden judgement, negotiation, supervision, and authority. This context illustrates how ADI can increase without requiring a formal transfer of final decision authority.

Second, school leadership and AI integration illustrate how technological capability interacts with organisational conditions before acquiring substantive institutional influence. Evidence from school leaders indicates that AI integration is shaped by leadership perceptions, organisational readiness, infrastructure, professional capability, and institutional conditions (Marrone et al., 2025). This context demonstrates why comparable AI capabilities may produce different levels of operational embedding and actual decision influence across educational organisations.

Third, institutional AI governance in higher education illustrates governance adaptation around expanding AI use. University policy research documents institutional attempts to define responsibilities, permissible practices, and governance arrangements for generative AI (An et al., 2025). Broader systematic and international evidence identifies expanding AI involvement in institutional governance alongside continuing requirements for human oversight, accountability, policy adaptation, and institutional support (Crompton et al., 2026; Jiang & Abdullah, 2026). This context illustrates the EGC and GRC components of the model by showing why governance arrangements require continuing reassessment as institutional AI use develops.

Together, these documented contexts establish the analytical applicability of the educational AI-GCM extension across administrative decision-making, school leadership, and higher-education governance. They serve as institutional illustrations rather than empirical tests or validation of AI-GCM.

DISCUSSION

Answer to the Research Question

The theoretical synthesis indicates that changes in AI capability generate governance–capability misalignment not through technological advancement alone, but when those changes alter actual AI decision influence without corresponding adjustment in educational governance configuration. Educational institutions may expand AI functionality, deploy AI across additional activities, or increase the sophistication of AI-supported processes while retaining adequate governance. The critical condition is deterioration in the substantive correspondence between the AI configuration operating within institutional decisions and the governance configuration intended to authorise, oversee, constrain, intervene in, and assign accountability for those decisions.

The central relationship can therefore be expressed as:

Material Change in AI Capability
→ Change in Actual AI Decision Influence

-

Insufficient Adjustment of Educational Governance Configuration
→ Potential Governance–Capability Misalignment

This relationship answers the first component of the research question by locating misalignment in changing correspondence rather than AI adoption itself. Educational research demonstrates expanding AI involvement in administrative and governance processes (Jiang & Abdullah, 2026), while educational leadership research indicates that AI can alter decision environments without eliminating the formal authority and responsibilities of educational leaders (Chen & Ma, 2026; Dai et al., 2025).

The consequences follow from this deteriorating correspondence. Formally assigned leadership authority may diverge from substantive AI influence; procedural human participation may persist while meaningful intervention capacity weakens; and institutional accountability may remain attached to actors whose substantive control over AI-mediated decision formation has diminished. These consequences are consistent with meaningful human control research showing that formal human participation cannot substitute for substantive authority and practical intervention capacity (Cavalcante Siebert et al., 2023).

The second component of the research question concerns restoration of alignment. Within the educational extension, Governance Recalibration Capacity provides the corrective mechanism. Recalibration requires institutions to detect governance-relevant changes in capability or influence, assess their implications, authorise a response, and adjust decision rights, oversight, intervention,

escalation, traceability, or accountability where necessary. Restored alignment is therefore produced through governance reconfiguration rather than through recognition alone and remains potentially temporary as AI-mediated decision systems continue to evolve.

Theoretical Contribution

The theoretical contribution of this study lies in the extension and specification of AI-GCM within educational leadership and AI-mediated institutional decision-making. It specifies how the parent theory's dynamic governance–capability relationship operates within educational institutions by connecting changing AI capability and actual decision influence with educational governance configuration, leadership consequences, and governance recalibration.

This specification integrates research streams that address related but analytically distinct components of the governance problem. Educational AI governance research establishes institutional policies, accountability arrangements, oversight requirements, and adaptive governance processes (Crompton et al., 2026; Jiang & Abdullah, 2026). Human–AI decision research examines differentiated roles for AI systems and educational leaders (Dai et al., 2025), while meaningful human control research establishes that substantive control depends on authority and practical intervention capacity rather than formal human presence alone (Cavalcante Siebert et al., 2023).

The contribution lies in relating these components through a dynamic governance-fit mechanism. Four distinctions are central:

Capability ≠ Authority
Formal Authority ≠ Actual Decision Influence
Human Participation ≠ Meaningful Human Control
Governance Adequacy at t_0 ≠ Governance Adequacy at t_1

The final distinction provides the temporal foundation of the extension. Governance arrangements may be substantively appropriate when established yet become progressively less appropriate as AI capability, operational embedding, or actual decision influence changes. Governance adequacy must therefore be evaluated through continuing correspondence rather than inferred from the continuing existence of governance structures.

The resulting contribution is an explanatory model of how educational AI governance can become misaligned despite the continued formal presence of governance structures, and how educational leadership can restore substantive governance–capability correspondence as AI-mediated decision systems evolve.

Relationship to Existing Explanations and the Parent Theory

The educational extension is positioned within, rather than against, existing explanations. AI-GCM provides the parent theoretical architecture for explaining how initially substantive governance can lose fit as AI capability and organisational use evolve, including divergence, organisational lag, governance reconfiguration, temporary fit, and organisational learning (Alaali, 2026b). The present study specifies how this broader process operates through educational leadership authority, actual AI decision influence, meaningful human control, and institutional accountability.

Educational AI governance research already establishes responsible-AI principles, institutional policies, governance structures, monitoring, accountability, and adaptive governance arrangements (Crompton et al., 2026; Jiang & Abdullah, 2026). The educational extension connects these arrangements to the continuing question of whether governance remains substantively aligned with changing AI capability and decision influence.

Educational leadership research identifies changing leadership roles, competencies, responsibilities, and institutional adaptation associated with AI (Chen & Ma, 2026). The extension adds a relational and temporal question: whether formally retained leadership authority continues to correspond to substantive influence over AI-mediated decision formation as institutional AI use changes.

Human–AI decision-making research establishes differentiated and collaborative roles for humans and AI systems (Dai et al., 2025). The extension separates formal role allocation from actual AI decision influence and examines whether the correspondence between them changes as AI becomes more deeply embedded in institutional decision processes.

Meaningful human control research establishes conditions under which human involvement constitutes substantive control rather than procedural presence (Cavalcante Siebert et al., 2023; Veluwenkamp, 2022). The educational extension situates these conditions within an evolving governance configuration and explains how substantive control may weaken even while formal human participation remains.

Implications for Educational Leadership

The educational extension shifts leadership attention from the mere presence of AI governance toward the continuing fit between governance arrangements and actual AI use. Educational leaders should therefore distinguish between:

What is the AI system formally permitted to do?
And **how much substantive influence is the AI system actually exercising over institutional decisions?**

A difference between these conditions does not itself establish governance–capability misalignment. It provides a governance-relevant signal that existing decision rights, oversight requirements, intervention mechanisms, escalation arrangements, traceability, and accountability allocations may require reassessment.

Governance review may therefore be triggered by changes across four categories. Technical triggers include material changes in AI functionality and significant vendor or system changes. Operational triggers include deployment into additional institutional functions and increased routine reliance on AI recommendations. Decision-risk triggers include movement into higher-consequence or more difficult-to-reverse decisions. Governance triggers include declining practical capacity for human intervention and changes in institutional accountability exposure. These triggers are consistent with educational evidence emphasising leadership capability, organisational readiness, institutional support, policy adaptation, and continuing governance as AI adoption develops (An et al., 2025; Chen & Ma, 2026; Crompton et al., 2026; Kafa, 2025; Marrone et al., 2025).

The practical implication is a shift toward capability-triggered and influence-sensitive governance review. Governance should not depend solely on fixed review schedules or the conditions under which an AI system was originally authorised. Material changes in technical capability, operational deployment, decision risk, or governance conditions provide grounds for reassessing whether the current governance configuration still provides substantive leadership authority, meaningful oversight, practical intervention capacity, and appropriately aligned accountability.

Such reassessment should focus on governance mechanisms that educational institutions can actually control: AI-use permissions, decision rights, mandatory review, acceptance or rejection authority, escalation, audit logging, procurement conditions, deployment restrictions, vendor obligations, and suspension authority. This places educational leadership responsibility at the level of institutional

governance rather than assuming control over proprietary AI architecture or vendor-controlled technical systems.

FUTURE EMPIRICAL VALIDATION

Construct Operationalisation

The theoretical relationships developed in this study require subsequent empirical validation. Future research should operationalise the five principal constructs: AI Capability Configuration (ACC), Actual AI Decision Influence (ADI), Educational Governance Configuration (EGC), Governance–Capability Misalignment (GCM), and Governance Recalibration Capacity (GRC). Measurement should preserve their distinct theoretical roles rather than treating them as interchangeable dimensions of AI governance. In particular, empirical designs should distinguish what an AI system is capable of doing from the substantive influence its outputs exercise and distinguish the presence of governance arrangements from their continuing correspondence to changing AI capability and influence.

ADI presents a particularly important measurement challenge because formal decision authority may remain human while AI outputs substantially shape decision formation. Potential indicators include reliance on AI recommendations, influence over prioritisation and available alternatives, frequency of acceptance or challenge, and the practical capacity of authorised human actors to modify, reject, or escalate AI-mediated decisions. Research on AI-assisted educational decision-making and meaningful human control provides a foundation for developing such measures (Cavalcante Siebert et al., 2023; Dai et al., 2025).

Validation Sequence

A staged empirical programme is appropriate. Stage 1 — Content Validation should use expert assessment to evaluate construct definitions, domain coverage, item relevance, and conceptual separation. Stage 2 — Measurement Validation should develop and test measurement items, including reliability, convergent validity, discriminant validity, and overall construct validity.

Stage 3 — Institutional Testing should examine P1–P7 across educational institutions and decision domains. Initial cross-sectional designs could assess relationships among ACC, ADI, EGC, GCM, and GRC and compare institutions characterised by different levels of AI adoption, governance maturity, and decision exposure.

Stage 4 — Longitudinal Validation is particularly important because the proposed mechanism explicitly concerns change over time:

Capability Change → Decision Influence Change → Governance Response/Non-response → Misalignment → Recalibration → Governance Outcome

Longitudinal evidence would permit examination of whether changes in AI capability and influence precede deterioration in governance fit and whether subsequent recalibration reduces persistent misalignment. This design is consistent with educational research emphasising evolving AI deployment, changing leadership demands, and continuing institutional governance adaptation (Chen & Ma, 2026; Crompton et al., 2026; Jiang & Abdullah, 2026).

Potential Human and Organisational Conditions

Several variables remain deliberately outside the core theoretical architecture but provide directions for empirical testing. At the individual level, AI literacy, trust in AI, and leadership AI expertise may influence how decision-makers interpret and rely on AI outputs. At the organisational level, staff reliance on AI, institutional dependence on AI-supported processes, and the availability of viable human

alternatives may affect the translation of AI capability into actual decision influence (Chen & Ma, 2026; Kafa, 2025).

These factors can be examined as antecedents, moderators, or contextual conditions without expanding the core theoretical architecture.

Alternative Explanations

Empirical validation should distinguish GCM from competing explanations. Apparent governance problems may result from generally weak institutional governance, inadequate AI literacy, resource constraints, implementation failure, leadership resistance, vendor dependence, or regulatory uncertainty rather than changing governance–capability correspondence. Educational research already identifies leadership capability, institutional readiness, resources, and organisational support as important conditions affecting AI implementation (Crompton et al., 2026; Jiang & Abdullah, 2026; Kafa, 2025).

Future tests should therefore determine whether observed deterioration in governance fit follows material change in AI capability or actual decision influence rather than merely reflecting governance weakness that existed from the outset.

LIMITATIONS

This study has five principal limitations. First, the educational extension of AI-GCM remains conceptual. The relationships represented in the model and Propositions P1–P7 are theoretically derived from an integrative synthesis and require empirical testing before causal or predictive claims.

Second, the principal constructs require further measurement development. ACC, ADI, EGC, GCM, and GRC perform distinct theoretical functions, but validated measures capable of separating these functions within educational institutions have not yet been established.

Third, Actual AI Decision Influence may be difficult to observe empirically. Formal decision rights are often documented, whereas substantive influence may operate through recommendations, prioritisation, routinised reliance, default options, or other less visible mechanisms. Distinguishing ADI from ordinary decision support and voluntary human reliance may therefore require multiple forms of evidence.

Fourth, the model is subject to institutional and jurisdictional variation. Educational governance structures, leadership authority, regulatory requirements, accountability arrangements, and decision rights differ across schools, universities, educational systems, and jurisdictions. The relative strength and operation of the proposed relationships may consequently vary across contexts (Crompton et al., 2026; Jiang & Abdullah, 2026).

Finally, technological and vendor boundaries may constrain institutional governance capacity. Educational institutions using externally controlled AI systems may have limited visibility into proprietary architecture, training data, model weights, or technical changes. Governance recalibration may therefore depend substantially on institutionally controllable mechanisms such as deployment permissions, human review, escalation, procurement requirements, vendor obligations, and suspension authority rather than direct technical modification.

CONCLUSION

This study examined how changes in AI capability and actual AI decision-making influence can generate governance–capability misalignment within educational institutions and how educational leadership can respond when such misalignment emerges. Drawing on educational AI governance, educational

leadership, human–AI decision-making, and meaningful human control research, the study specified the following theoretical sequence:

AI Capability Change
→ Operational Embedding and Capability-to-Influence Translation
→ Actual AI Decision Influence
→ Persistence of Existing Educational Governance Configuration
→ Potential Governance–Capability Misalignment
→ Leadership Authority / Human Control / Accountability Consequences
→ Governance Recalibration
→ Potential Restored Governance–Capability Fit

Increasing AI capability does not itself constitute governance failure. Misalignment becomes possible when changes in capability acquire substantive decision influence without corresponding adjustment in the governance arrangements through which educational institutions allocate authority, oversight, intervention, escalation, traceability, and accountability. Governance adequacy therefore depends on continuing correspondence between these arrangements and the AI-mediated decision environment rather than on the formal presence of governance structures alone.

By extending AI-GCM into educational leadership, the study connects AI Capability Configuration, Actual AI Decision Influence, Educational Governance Configuration, Governance–Capability Misalignment, and Governance Recalibration Capacity within a dynamic explanatory architecture. The extension distinguishes technological capability from decision influence, formal authority from substantive influence, human participation from meaningful human control, and governance presence from continuing governance adequacy.

The resulting analytical question is not simply whether educational institutions possess AI governance, but whether that governance continues to fit the AI-mediated decision environment it is intended to govern. As AI capability, operational embedding, and institutional reliance evolve, maintaining substantive fit requires educational leadership capable of recognising governance-relevant change and translating it into timely governance reassessment and, where necessary, institutional recalibration.

REFERENCES

Alaali, H. M. H. (2025a). *DFAS-EEP: Editorial ethics protocol for governance integrity in scientific publishing*. SSRN. https://doi.org/10.2139/ssrn.5375339

Alaali, H. M. H. (2025b). The DFAS-FEP Protocol: A global governance standard for responsible AI use and authorship integrity in financial modelling. SSRN. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5260517

Alaali, H. M. H. (2026a). AI governance in financial systems: A PGCC and emerging markets convergence framework for accountability, sovereignty, and operational control. *Journal of Countries Studies*, 1–58. https://doi.org/10.22059/JCOUNTST.2026.412773.1476

Alaali, H. M. H. (2026b). *When AI outruns its governance: A dynamic capabilities theory of governance–capability misalignment and reconfiguration.* Zenodo. https://zenodo.org/records/21986921

An, Y., Yu, J. H., & James, S. (2025). Investigating the higher education institutions' guidelines and policies regarding the use of generative AI in teaching, learning, research, and administration. *International Journal of Educational Technology in Higher Education*, *22*, Article 10. https://doi.org/10.1186/s41239-025-00507-3

Cavalcante Siebert, L., Lupetti, M. L., Aizenberg, E., Beckers, N. W. M., Zgonnikov, A., Veluwenkamp, H. M., Abbink, D. A., Giaccardi, E., Houben, G. J. P. M., Jonker, C. M., van den Hoven, M.

J., Forster, D., & Lagendijk, R. L. (2023). Meaningful human control: Actionable properties for AI system development. *AI and Ethics, 3*(1), 241–255. https://doi.org/10.1007/s43681-022-00167-3

Chan, C. K. Y. (2023). A comprehensive AI policy education framework for university teaching and learning. *International Journal of Educational Technology in Higher Education, 20*, Article 38. https://doi.org/10.1186/s41239-023-00408-3

Chen, L. H., & Ma, M. (2026). Leading in the AI age: A systematic review of global perspectives on AI and educational leadership. *ECNU Review of Education, 9*(2), 1–19. https://doi.org/10.1177/20965311261446186

Crompton, H., Burke, D., Nickel, C., Bozkurt, A., Miao, F., Sharples, M., Greene, J. A., Parsons, D., Gill-Simmen, L., Edmett, A., Pegrum, M., de Waard, I., Bonk, C. J., Garcia, M. B., Curry, J. H., Lindsey, L. A., Yang, M., Marshall, S., Bali, M., . . . Yu, S. (2026). Governing generative AI in higher education: A global Delphi study on policy and practice. *International Journal of Educational Technology in Higher Education, 23*(1), Article 21. https://doi.org/10.1186/s41239-026-00602-z

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*, 1471–1487. https://doi.org/10.1007/s13384-024-00771-8

Enqvist, L. (2023). 'Human oversight' in the EU artificial intelligence act: What, when and by whom? *Law, Innovation and Technology, 15*(2), 508–535. https://doi.org/10.1080/17579961.2023.2245683

Jiang, X., & Abdullah, Z. (2026). AI-enabled governance in higher education: A systematic review of applications, outcomes, and emerging implications. *Frontiers in Education, 11*, Article 1856440. https://doi.org/10.3389/feduc.2026.1856440

Johari, A. M. Y., Ghani, M. F. A., Radzi, N. M., Ayub, A., & Ghani, M. A. (2023). Challenges of school leaders' digital leadership: An initial study in Bintulu, Sarawak, Malaysia. *Jurnal Akuntabilitas Manajemen Pendidikan, 11*(1), 57–65. https://doi.org/10.21831/jamp.v11i1.58443

Kafa, A. (2025). Exploring integration aspects of school leadership in the context of digitalization and artificial intelligence. *International Journal of Educational Management, 39*(8), 98–115. https://doi.org/10.1108/IJEM-11-2024-0703

Lazcoz, G., & de Hert, P. (2023). Humans in the GDPR and AIA governance of automated and algorithmic systems: Essential pre-requisites against abdicating responsibilities. *Computer Law & Security Review, 50*, Article 105833. https://doi.org/10.1016/j.clsr.2023.105833

Marrone, R., Fowler, S., Bathakur, A., Dawson, S., Siemens, G., & Singh, C. (2025). Perceptions and perspectives of Australian school leaders on the integration of artificial intelligence in schools. *School Leadership & Management, 45*(1), 30–52. https://doi.org/10.1080/13632434.2024.2425019

Matthias, A. (2004). The responsibility gap: Ascribing responsibility for the actions of learning automata. *Ethics and Information Technology, 6*(3), 175–183. https://doi.org/10.1007/s10676-004-3422-1

Polat, M., Karataş, İ. H., & Varol, N. (2025). Ethical artificial intelligence (AI) in educational leadership: Literature review and bibliometric analysis. *Leadership and Policy in Schools, 24*(1), 46–76. https://doi.org/10.1080/15700763.2024.2412204

Sposato, M. (2025). Artificial intelligence in educational leadership: A comprehensive taxonomy and future directions. *International Journal of Educational Technology in Higher Education, 22*, Article 20. https://doi.org/10.1186/s41239-025-00517-1

Veluwenkamp, H. M. (2022). Reasons for meaningful human control. *Ethics and Information Technology, 24*(4), Article 51. https://doi.org/10.1007/s10676-022-09673-8