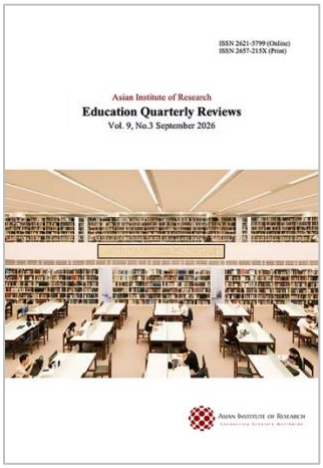

Education Quarterly Reviews

Chin, P. P. L. (2026). Which Decisions Remain the Educator's Own: The Hybrid Decision-Making Model for AI-Integrated Education. *Education Quarterly Reviews*, 9(3), 78-96.

ISSN 2621-5799

DOI: 10.31014/aior.1993.09.03.726

The online version of this article can be found at:
https://www.asianinstituteofresearch.org/

Published by:
The Asian Institute of Research

The *Education Quarterly Reviews* is an Open Access publication. It may be read, copied, and distributed free of charge according to the conditions of the Creative Commons Attribution 4.0 International license.

The Asian Institute of Research *Education Quarterly Reviews* is a peer-reviewed International Journal. The journal covers scholarly articles in the fields of education, linguistics, literature, educational theory, research and methodologies, curriculum, elementary and secondary education, higher education, foreign language education, teaching and learning, teacher education, education of special groups, and other fields of study related to education. As the journal is Open Access, it ensures high visibility and increases citations for all research articles published. The *Education Quarterly Reviews* aims to facilitate scholarly work on recent theoretical and practical aspects of education.

ASIAN INSTITUTE OF RESEARCH
Connecting Scholars Worldwide

The Asian Institute of Research
Education Quarterly Reviews
Vol.9, No.3, 2026: 78-96
ISSN 2621-5799
Copyright © The Author(s). All Rights Reserved
DOI: 10.31014/aior.1993.09.03.726

Which Decisions Remain the Educator's Own: The Hybrid Decision-Making Model for AI-Integrated Education

Pauline P. L. Chin[1]

[1] Meragang Sixth Form College
ORCID ID: https://orcid.org/0000-0001-5635-9141

Abstract
When educators working with artificial intelligence (AI) are asked which decisions remain their own to make, most know the boundary exists. Few have a structure for it. This paper addresses that absence directly by proposing the Hybrid Decision-Making Model, a three-tier framework that assigns decision authority across AI-integrated higher education practice. Tier 1 covers routine, reversible decisions where AI leads and the human retains oversight. Tier 2 covers decisions requiring both AI-provided data and human contextual judgement, where neither contribution alone produces a valid outcome. Tier 3 covers decisions whose irreversible consequences, ethical weight, and relational specificity make human authority a structural requirement rather than a preference. The HDM is grounded in three theoretical foundations: Simon's (1947, 1990) account of bounded rationality, which explains why the absence of decision authority produces predictable drift towards AI-generated outputs; Engelbart's (1962) and Baker's (2016) principle of augmentation, which establishes that AI should extend human capability rather than replace it; and the hybrid intelligence and governance frameworks of Holstein et al. (2019, 2020), Akata et al. (2020), and IMDA and PDPC (2020), which provide the dimensions and human involvement taxonomy from which the three tiers are derived. Empirical support is drawn from a qualitative study of seven higher education educators who consistently drew the same structural distinctions the HDM formalises, in the absence of a formally named decision authority structure for AI-integrated practice. The model does not prescribe institutional behaviour. It names the conditions under which decision authority is maintained. The structure was already there. This paper simply gives it a name.

Keywords: Artificial Intelligence, Higher Education, Decision Authority, Human-AI Collaboration, Hybrid Decision-Making, AI Governance

1. Introduction

A science educator, when asked about artificial intelligence in the classroom, did not question whether the technology belonged there. The question that arose was more precise: which decisions remained the educator's own to make. This finding, drawn from a qualitative study of seven higher education educators (Chin, 2025), points to a structural problem that enthusiasm for AI integration has so far outpaced. A review of 24 national AI policy strategies found that the use of AI in education is largely absent from policy conversations, whilst the instrumental value of education in producing an AI-ready workforce is overwhelmingly prioritised (Schiff, 2022). Most educators working with AI today would recognise this clearly: they know the boundary between human and machine decision-making exists, they feel it every time a choice must be made, but they do not yet have a map for it. Higher education institutions are therefore adopting AI tools without an agreed structure for determining which

decisions AI should lead, which educators should lead, and which require the contribution of both. Even national governance frameworks that have begun to address AI in organisational settings have not resolved this question at the level of classroom practice (AITI, 2025; IMDA and PDPC, 2020). The problem is not one of willingness. It is one of design.

The consequences of this absence are already visible, though rarely named. When institutional guidance on artificial intelligence is unclear, educators do not stop; they adapt, drawing their own boundaries as they go. This produces inconsistency across classrooms (Chan, 2023). Without a shared structure, decisions made under time pressure tend naturally towards convenience. Pedagogical judgement shifts gradually as a result: the question moves from what does this student need to what does the AI suggest, and the two are not always the same. Students, meanwhile, receive different messages about acceptable use depending on which classroom they are in, which produces uneven habits and an uneven relationship with their own thinking. Uncertainty about academic integrity follows, not from dishonesty on anyone's part, but from the absence of clear and shared expectations. Moments of genuine good practice occur across institutions, but without a common structure they remain individual rather than collective achievements. The difficulty, therefore, is not that educators or institutions have failed. It is that the situation has moved faster than the frameworks available to manage it.

The research field has made substantial progress in identifying the tensions that AI integration introduces into higher education. Studies have documented the risk that over-reliance on AI-generated material weakens students' capacity for independent thinking (Chin, 2025; Popenici and Kerr, 2017), that academic integrity becomes uncertain when the authorship of student work is unclear (Chin, 2025; Chan, 2023), and that data privacy and ethical responsibility remain inadequately resolved in most institutional contexts (Chan, 2023; Floridi and Cowls, 2019; Miao and Holmes, 2021). The potential for AI systems to produce biased or inaccurate outputs has been noted as a concern that existing safeguards have not yet addressed sufficiently (Chan, 2023), whilst unequal access to AI tools risks widening educational disparities that institutions are already working to close (Schiff, 2022; Miao and Holmes, 2021). The governance challenges that arise when AI is integrated into institutional decision-making processes have been documented in regional policy contexts, including within the Southeast Asian setting in which this paper is situated (Yar et al., 2024). The field has also examined educator attitudes, student behaviour, and policy developments in AI adoption with increasing depth and careful development (Chin, 2025; Chan, 2023; Schiff, 2022). A scoping review of 32 empirical studies confirmed that balancing AI-assisted and human-centred assessment remains an unresolved challenge, with educators calling for clearer boundaries around what AI can and cannot determine in student work (Xia et al., 2024). In this sense, the problem is well understood. What the literature has not yet produced is a practical structure for decision-making: a clear account of which decisions artificial intelligence may lead, which must remain under human judgement, and which should involve both.

That absence has a common source. Each risk identified in the literature points to the same underlying issue: the absence of clear control over decision-making in AI-integrated practice. The risks include weakened independent thinking (Chin, 2025; Popenici & Kerr, 2017), uncertain academic integrity (Chin, 2025; Chan, 2023), unresolved ethical responsibility (Chan, 2023; Floridi & Cowls, 2019), unchecked AI outputs (Chan, 2023), and unequal access (Schiff, 2022; Miao & Holmes, 2021). In each case, the difficulty arises not from the presence of artificial intelligence itself, but from the absence of guidance on who should decide, when that decision should be made, and on what basis.

It is the quiet centre of all the literature has been circling. The field has mapped where control is missing, even when it has not named the mapping as such. This is not simply a practical difficulty. Simon (1947, 1990) demonstrated that human decision-making operates under real cognitive constraints: individuals do not optimise across all available choices but settle for what is sufficient given the time, information, and capacity available. Without a deliberate structure to direct judgement, decisions made under pressure default to convenience rather than principle. The absence of control is not merely inconvenient. Simon (1947, 1990) shows it is structurally dangerous.

Three bodies of work established the conditions that made a response to this danger both necessary and possible. Baker (2016) rejected full automation, demonstrating that AI should extend human capability rather than replace

it and that the human must remain present in decision-making. Holstein et al. (2019, 2020) established that human-AI collaboration is not natural or automatic. It must be deliberately designed, and roles cannot be left vague. Akata et al. (2020) identified the distribution of decision-making between human and machine as the central unresolved problem in hybrid intelligence research. Each contribution advanced the field. None produced the structure itself. Each described how work should be distributed between human and machine. None specified who holds authority when a judgement must be made.

This study addresses that directly. Drawing on semi-structured interviews with seven experienced higher education educators, it examines how these educators drew decision-making boundaries in their own teaching, in the absence of formal institutional guidance (Chin, 2025). The study is small and qualitative by design. Its value lies not in breadth but in the depth of seven educators' lived experiences, which reveal what large-scale surveys cannot: the precise distinctions practitioners make when no framework exists to guide them. The central question the study asks is this: how should decision-making authority be distributed between artificial intelligence and human educators in higher education practice? What the study reveals is that educators were already making these distinctions in practice, intuitively and consistently, long before a formal structure existed to support them. The Hybrid Decision-Making Model does not invent a new structure. It names and formalises one that practitioners have already been living.

The paper proceeds as follows. Section 2 reviews the three bodies of literature from which the HDM is constructed: hybrid intelligence frameworks, teacher-AI collaboration research, and the division of labour between human and machine. Section 3 establishes the theoretical foundations of the model. Section 4 presents the Hybrid Decision-Making Model in full. Section 5 draws on the empirical study to validate the framework across its three tiers. Section 6 offers reflections on how the model might inform institutional practice. Section 7 discusses the paper's contributions in relation to the field. Section 8 identifies directions for future research. The science educator whose question opened this paper asked which decisions remained their own to make. This paper answers that question. Not with a reassurance, but with a structure.

2. Literature Review

Three bodies of scholarship inform the theoretical foundations of the Hybrid Decision-Making Model. The first examines hybrid intelligence in education. The second examines teacher-AI collaboration. The third examines the division of labour between human and machine. Each advances the field substantially. Each stops at the same point: no study identified in this review specifies how decision-making authority should be distributed in practice.

2.1. Hybrid Intelligence in Education

Akata et al. (2020) established that combining human and artificial capability effectively requires systems that are collaborative, adaptive, responsible, and explainable, and identified the distribution of decision-making between human and machine as a problem the field had not yet resolved. Bredeweg and Kragten (2022) demonstrated through a case study in secondary education that intelligent tutoring systems work more effectively when teachers remain involved, not by doing what the system does alongside it, but by focusing on what it cannot do: interpreting data about student progress and making pedagogical judgements that require human experience.

Holstein et al. (2020) identified four dimensions structuring human and AI instruction: instructional goals, relevant information, instructional actions, and decision-making processes. Their contribution is showing that these dimensions exist and can be mapped. What the framework does not address is who holds authority at each dimension. It describes how human and AI instruction can be combined but does not specify which actor makes the final decision at each point, or on what basis.

2.2. Teacher-AI Collaboration

Baker (2016) argued that rather than building AI systems capable of replacing teacher decision-making, the field should focus on building systems that provide teachers with better data and analysis, leaving the decisions

themselves with the educator. Baker named this intelligence amplification. The human decides. The AI provides the information on which that decision is based.

Holstein et al. (2019) investigated what intelligence amplification looks like in practice through a study of teacher and student needs in AI-enhanced classrooms. They found that teachers wanted real-time data from AI systems to support their decisions, not automated decisions made on their behalf. The design of collaboration must be deliberate. What Holstein et al. (2019) do not specify is the decision authority structure that design should produce.

Paiva and Bittencourt (2020) confirmed that this collaborative model is achievable in real settings. Their T-Partner authoring tool processes educational data to support instructors in making pedagogical decisions without making those decisions for them. This confirms that the model Baker proposed holds under real working conditions.

2.3 The Division of Labour Between Human and Machine

Engelbart (1962) argued that computing technology exists to augment human intellect rather than automate it. That principle sits at the foundation of the HDM.

Chou, Huang and Lin (2011) applied this principle in a higher education tutoring context, dividing tutoring responsibilities between software and teacher. The software evaluated answers and generated hints. The teacher interpreted student thinking and made relational decisions. Their findings confirmed that this division reduced teacher workload without reducing student support quality. The main condition for success was that each party's responsibilities were defined clearly from the outset.

Holstein and Aleven (2022) confirmed the same pattern in K-12 classrooms through a study involving Lumilo, a tool displaying real-time student learning data on smart glasses worn by the teacher. Students learned more when the teacher used Lumilo alongside the AI tutoring system because the teacher remained the decision-maker: Lumilo provided the information and the teacher decided how to respond. None of these studies, however, specifies who holds decision authority when a judgement call is required. Allocating tasks is not the same as assigning decision authority. The HDM addresses that distinction.

2.4 The Common Gap: Distributing Work Without Distributing Decisions

There is a precise distinction between distributing work and distributing decisions. Distributing work assigns tasks. Distributing decisions assigns authority: when a judgement is required, whose call is it? A teacher can be assigned the task of reviewing AI-generated feedback without it being established whether that teacher has the authority to change it, question its basis, or set it aside. The task has been allocated. The authority has not.

The consequences of leaving decision authority unspecified are not theoretical. Simon (1947, 1990) demonstrated that without a structure that clearly assigns decision authority, an educator working through a full teaching day will not pause to consider whether a particular decision is theirs to make. The default will be convenience. Over time, convenience quietly replaces principle. Decisions drift towards the AI. Neither drift is dramatic. Neither is immediately visible. Both are consequential.

Akata et al. (2020) identified the distribution of decision-making as the central unresolved problem in hybrid intelligence research. Holstein et al. (2019) established that human-AI collaboration must be deliberately designed but did not specify what the design of decision authority should look like. Baker (2016) argued that the human must remain in the decision loop but did not define which decisions that loop should contain. It is the precise point at which the existing research reaches its limit.

A concurrent contribution moves further Into this territory than the frameworks above. Borchers, Viberg, and Kizilcec (2026) introduce the Agency Allocation Framework (AAF), which reframes learner agency as the allocation of decision-making authority across learners, educators, institutions, and AI systems in large-scale, automated learning environments. Their framework treats decision authority, rather than task distribution, as its

central object, and it names four recurring challenges that keep that authority from being studied or designed with precision: conceptual ambiguity, the difficulty of measuring agency at scale, the trade-off between agency and efficiency, and the redistribution of agency through AI mediation. Its diagnosis strengthens the case that a structural resolution is needed. Its locus and its purpose, however, differ from the one this paper proposes. The AAF addresses learner agency in large-scale automated systems, and it stops at naming the challenges that make authority hard to allocate. The HDM addresses educator agency in general AI-integrated higher education practice, and it does not stop at diagnosis. It proposes the tiers themselves, and a test for assigning a decision to one of them. Where the AAF clarifies why decision authority has remained elusive, the HDM specifies where it should sit.

The Hybrid Decision-Making Model begins where that limit ends.

3. Theoretical Foundations

The literature review established three findings that this section addresses directly. First, human-AI collaboration produces better outcomes when deliberately designed (Akata et al., 2020; Holstein et al., 2019). Second, existing frameworks have distributed work with precision but none specifies who holds authority at the moment of judgement (Baker, 2016; Holstein et al., 2020; Paiva and Bittencourt, 2020). Third, in the absence of a clear decision authority structure, cognitive pressure produces consistent and predictable drift (Simon, 1947, 1990). The HDM is the theoretical response to all three findings.

The section proceeds In three stages. Section 3.1 draws on Simon (1947, 1990) to establish why decision authority must be made explicit. Section 3.2 draws on Engelbart (1962) and Baker (2016) to establish augmentation as the governing principle. Section 3.3 derives the HDM's three tiers from Holstein et al. (2019, 2020), Akata et al. (2020), and the human involvement taxonomy of IMDA and PDPC (2020), subsequently adopted by Brunei Darussalam as its national guide on AI governance and ethics (AITI, 2025).

3.1. Bounded Rationality and the Design of Decision Authority

Drift, as used in this paper, refers to a directional and incremental shift away from deliberate human judgement and towards uncritical acceptance of AI-generated outputs. Under the conditions Simon (1947, 1990) identifies, it is predictable.

Simon's account of bounded rationality begins with a structural observation about the limits of human cognition. Decision-making proceeds not by surveying all available options and selecting the best one, but by finding an option that meets a minimum threshold and stopping. Simon called this satisficing: the decision that is good enough, given the information and time available, is the decision that is made. The constraint is cognitive: working memory is finite, attention is depletable, and the number of decisions competing for both across a teaching day is not.

Search is also costly. Reviewing an AI output more carefully, questioning its basis, comparing it against independent professional judgement all require time and resource that are not freely available. Bounded rationality names the limit on decision-making capacity. Search cost names why that limit is not overcome even when the educator is aware of it.

In an AI-integrated classroom without a clear decision authority structure, both mechanisms operate simultaneously. Before a substantive decision can be made, a prior question must first be resolved: whose decision is this? That meta-decision consumes resource before the substantive judgement begins. Under time pressure, it resolves towards the most immediately available option which, in an AI-integrated classroom, is consistently the AI's output.

The result is drift. The AI produces a satisfactory output; the absence of a structural signal that scrutiny is required increases the cost of providing it; the output is accepted; each acceptance recognises the next. The drift is gradual because no single acceptance constitutes a recognisable failure. It is directional because the asymmetry of

cognitive cost consistently favours the AI's output. It is unnoticed because each individual decision is locally defensible. What accumulates is not negligence. It is a structural realignment of where educational judgement is exercised, produced not by intention but by the absence of design.

Awareness alone cannot correct drift. Training addresses what educators know about the problem. It does not change the conditions that produce it. When decision authority is clearly assigned, the educator no longer needs to resolve the prior question of whose decision this is. Cognitive resource that was being spent on the meta-decision becomes available for decisions that genuinely require deliberate human judgement. Distributing work without distributing decisions leaves the meta-decision unresolved at every point of practice. Under the conditions Simon describes, it resolves towards convenience. That is the direction a decision authority structure is designed to interrupt.

3.2. Augmentation as the Governing Principle

If bounded rationality produces drift when decision authority is unclear, the governing principle for distributing that authority must reduce cognitive load without displacing human judgement. Only one arrangement satisfies both conditions: augmentation.

Engelbart (1962) argued that computing technology exists to extend human intellectual capability, not replace it. A technology that replaces human decision-making removes the contribution that gives the decision its validity in complex situations. A technology that extends it preserves that contribution whilst reducing the cognitive burden required to exercise it. Baker (2016) translated this into a concrete decision hierarchy: build systems that provide teachers with better information, leaving the decision with the educator. The human decides. The AI provides the basis on which that decision is made. Augmentation reduces the cognitive cost of deciding without transferring the authority to decide.

Without augmentation, efficiency and scrutiny compete directly, which is the condition Section 3.1 identified as productive of drift. Augmentation dissolves that tension by allowing AI to handle information-processing tasks that consume cognitive resource, whilst preserving the educator's authority over the judgements that require human experience and ethical responsibility. The educator is not burdened by what the AI can do well. The educator is freed for what only the educator can do.

This principle runs through every tier of the HDM. At the task level, augmentation is efficiency with oversight. At the judgement level, it is support without substitution.

3.3. From Hybrid Adaptivity to Decision Authority

The three tiers of the HDM were already latent in the existing literature. This subsection makes them explicit.
Holstein et al. (2020) identified four dimensions along which human and AI instruction can be combined: instructional goals, relevant information, instructional actions, and decision-making processes. The framework demonstrates that hybrid adaptivity is achievable. What it does not specify is where authority is held at the moment of judgement.

Akata et al. (2020) established four criteria for productive human-AI systems: collaborative, adaptive, responsible, and explainable. The HDM accepts those four and adds a fifth: clear. A system that does not specify who decides remains structurally incomplete. Without clarity on decision authority, the meta-decision burden Section 3.1 identified is reproduced at every point of practice.

Singapore's Model AI Governance Framework (IMDA and PDPC, 2020), subsequently adopted by Brunei Darussalam (AITI, 2025), distinguished three levels of human involvement in AI-augmented decision-making: the human makes the final decision informed by AI analysis; AI leads whilst the human monitors and retains authority to intervene; and AI operates autonomously with oversight maintained at the system level. These levels were

designed for organisational governance. Translating them into a decision authority structure for higher education classroom practice is what the HDM performs.

That translation requires one additional dimension the governance framework does not supply: decision consequence, referring to the degree to which a decision is reversible, ethically significant, and dependent on contextual and relational knowledge that only human experience provides. When the level of human involvement is mapped against the level of consequence, three distinct and stable configurations emerge: AI leads with human oversight; both contribute and neither decides alone; and human primacy is a structural requirement. Each resolves the meta-decision burden by specifying in advance whose decision this is.

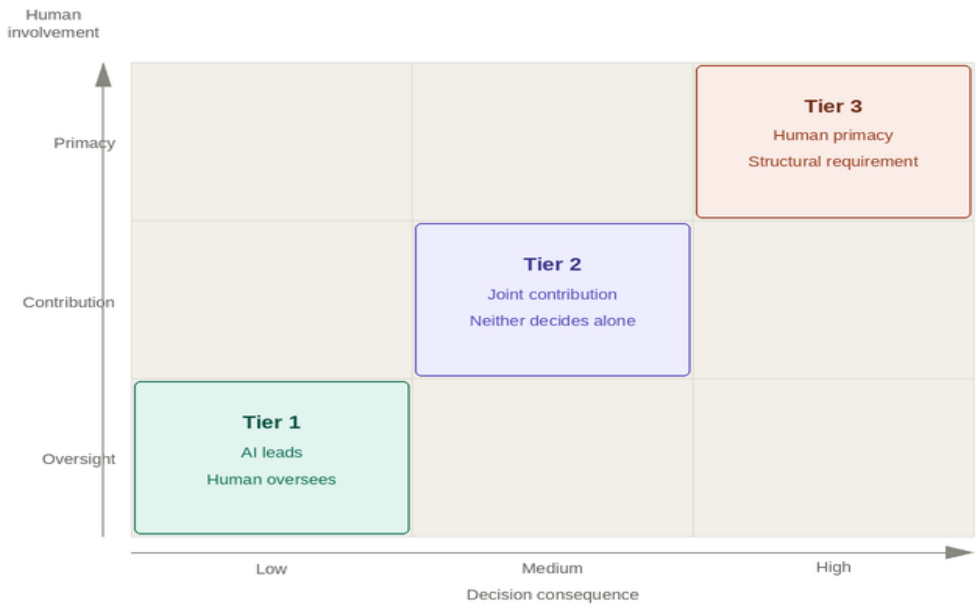

Figure 1: The interaction of human involvement and decision consequence producing three configurations of decision authority in the Hybrid Decision-Making Model.

The distinction between the HDM and its source frameworks is precise. Existing frameworks describe how humans and AI interact. The HDM specifies who holds authority when a judgement must be made. It is a decision authority taxonomy, assigning ownership of judgement across three levels of consequence. Without this resolution, existing frameworks remain descriptive rather than operational. The HDM replaces structural ambiguity with structural clarity. The premises established by Holstein et al. (2020), Akata et al. (2020), and IMDA and PDPC (2020) produce a condition that requires resolution. The HDM is that resolution.

Under conditions of bounded rationality, the absence of decision authority produces predictable drift. Augmentation provides the only governing principle capable of resolving that drift whilst preserving human judgement. When combined with the existing structure of hybrid intelligence and governance frameworks, a specific distribution of decision authority follows. What the field needed was not another account of how human and AI should work in concert, but a clear answer to who decides. What the literature has described as interaction, collaboration, and involvement is here resolved into authority.

4. The Hybrid Decision-Making Model

Section 3 established three findings in sequence: that the absence of decision authority produces predictable drift, that augmentation is the only principle capable of resolving that drift, and that applying that principle to the existing hybrid intelligence and governance frameworks produces a specific distribution of decision authority. This section presents that distribution. It is the Hybrid Decision-Making Model.

The HDM is not a set of options from which institutions select according to preference. It is a structure that locates every decision in AI-integrated higher education practice within one of three defined tiers. The tiers do not classify activities or workflows. They assign authority: who holds the right to make the final judgement, and on what basis. A decision is not negotiated into a tier. It belongs there because of what it is. Tier assignment is a structural determination based on the nature of the decision. Efficiency is a consequence of correct assignment, not a basis for it. An institution that assigns a decision to a tier because it is convenient to do so has not applied the HDM. It has misapplied it.

Tier assignment is determined by two criteria from Section 3.3: the consequence attached to the decision and the contextual and relational knowledge it requires. Both criteria are stable. What changes across institutional contexts is not the criteria but their application. A decision that is routine in one context may carry higher consequence in another and will be assigned to a higher tier accordingly. That is principled application of stable criteria, not situational flexibility. As AI capability and institutional context evolve, tier assignments should be reviewed. The criteria for assignment do not change.

The three tiers are related in three distinct ways. They are hierarchical in terms of consequence and human authority: Tier 3 carries more consequence and requires more human authority than Tier 1. They are complementary in terms of function: each tier handles a category of decisions that the other two cannot handle well, and collectively the three tiers cover the full decision environment of AI-integrated higher education practice. They are simultaneous in terms of operation: a teaching day does not proceed through Tier 1 decisions, then Tier 2, then Tier 3 in sequence. All three tiers operate concurrently across the full range of decisions an educator faces.

The pyramid structure (Figure 2) reflects this logic directly and its shape is not arbitrary. The base is wide because Tier 1 decisions are numerous: in any AI-integrated educational environment, the majority of decisions that AI can assist with are routine, high-volume, and low-consequence. The middle band is narrower because Tier 2 decisions require the specific condition of necessary integration, which applies to a smaller range of decisions than routine processing but a larger range than those carrying irreversible ethical consequence. The apex is narrow because Tier 3 decisions are the fewest in number and the highest in consequence: their specificity, ethical weight, and relational complexity mean that they cannot be routinised, scaled, or delegated.

Before the three subsections begin, a practical tier assignment test follows from the two criteria directly. When a decision must be placed, three questions apply in sequence. First: if an error in this decision is not corrected, can the student recover from its consequences without lasting harm, and does the decision call for no interpretive judgement that only the educator can provide? If yes: Tier 1. Second: does this decision require both AI-provided data and the educator's contextual knowledge to be formed well, such that neither contribution alone produces a valid outcome? If yes: Tier 2. Third: does the validity of this decision depend on human judgement in a way that cannot be delegated, because its consequences are irreversible, its ethical weight is significant, or its relational specificity exceeds what data can capture? If yes: Tier 3. When a decision sits at the boundary between two tiers, assigning too low risks structural drift. The HDM is designed to prevent that. The examples across the three subsections are drawn primarily from assessment-adjacent contexts for illustrative clarity. The HDM applies across the full range of educational decision-making in AI-integrated practice.

The three tiers are presented in full in the subsections that follow. Figure 2 provides a structural overview before the subsections begin.

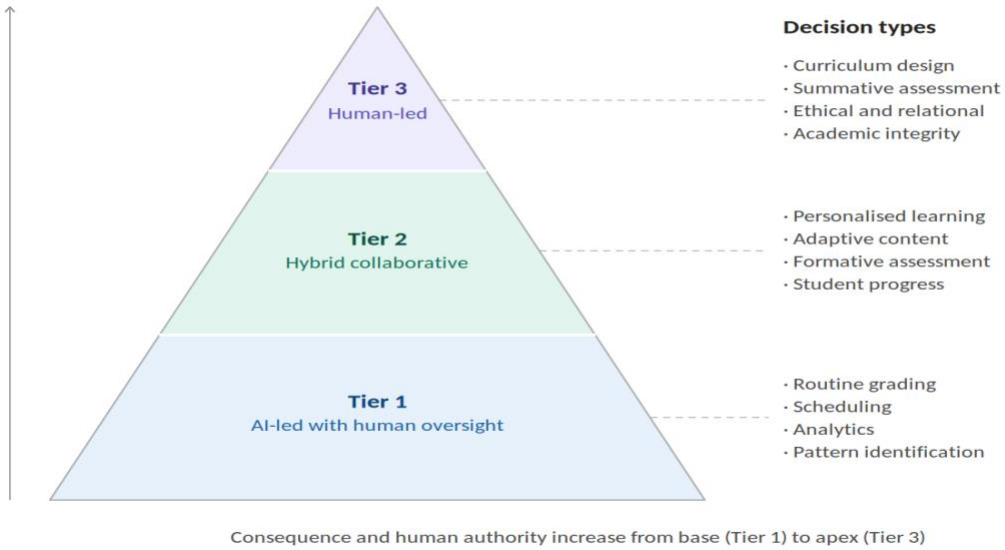

Figure 2: The Hybrid Decision-Making Model: distribution of decision authority across three tiers of AI-integrated higher education practice.

Tier 1 (base) covers routine, reversible decisions where AI leads and the human retains oversight. Tier 2 (middle) covers decisions requiring both data and contextual judgement, where the human holds final authority. Tier 3 (apex) covers decisions of high consequence and ethical weight, where human primacy is a structural requirement. Decision consequence and human authority increase from base to apex. Decision types listed are illustrative, not exhaustive.

4.1. Tier 1: AI-Led with Human Oversight

Tier 1 covers decisions that are routine, repeatable, and low in consequence. Human oversight at this tier does not mean passive monitoring. It means retained authority: the educator's unconditional right to review, question, and overrule any AI output.

Low consequence, as used here, refers to the reversibility of the outcome: where an error is recoverable without lasting harm to the student, and where correcting it calls for no interpretive judgement that only the educator can provide, the decision falls within Tier 1. A decision is not Tier 1 because it is convenient to automate. It is Tier 1 because its consequence is genuinely recoverable and its assessment requires no contextual or relational knowledge that only the educator holds. Where decisions of this nature recur at high volume and follow predictable patterns, AI is better positioned to execute them efficiently and consistently. Efficiency and consistency are the appropriate standards here, because these are decisions that should not vary according to who performs them. Routine grading of objective assessments, scheduling, analytics, and pattern identification share these properties. They are listed as illustrations of a category, not as a complete inventory. This configuration is not a design preference. It is the direct expression of the human-over-the-loop arrangement identified by IMDA and PDPC (2020), applied to the category of decisions where consequence is lowest and AI capability is highest.

The governing principle of this tier is clear: AI performs the task, but the human retains the right to decide whether the outcome stands. That is what separates Tier 1 from AI replacement. The task has been delegated. The authority has not. What keeps authority with the human is structural assignment, not personal vigilance. Because the decision has been categorised as Tier 1 in advance, the structure signals that every AI output is subject to human review.

The frequency with which that authority is exercised does not determine where it lies. An educator who reviews one in twenty AI-generated outputs has not transferred authority for the other nineteen. Authority is held structurally. Its exercise is discretionary. Its presence is unconditional. Most educators working with AI tools will

recognise this intuitively. The question is not whether they trust the output. It is whether they know they can change it. The HDM does not specify a minimum review threshold, because doing so would turn a decision authority framework into a compliance checklist. The model establishes where authority resides. Institutional practice determines how it is used.

This arrangement addresses the meta-decision burden identified in Section 3.1. At Tier 1, the question of whose decision this is has already been answered by the structure. The educator receives AI output already framed as subject to review. That framing removes the meta-decision, and in doing so interrupts the first and most pervasive condition under which drift takes hold.

Tier 1 expresses the augmentation principle established in Section 3.2 as efficiency with oversight. AI leading routine, high-volume tasks does not displace the educator. It frees attention for the decisions in Tier 2 and Tier 3 that require interpretation, ethical judgement, and relational knowledge that AI cannot provide. The HDM does not prevent educators from engaging closely with Tier 1 outputs. An educator who reviews AI-generated scores in detail retains full access to the data they contain. The base of the pyramid (Figure 2) is wide because Tier 1 decisions are numerous, and managing them through AI leadership with human authority retained is what makes the higher tiers of the model workable in practice. The structural safeguard described here operates at the level of individual decision-making. Whether institutions maintain the conditions under which that safeguard remains effective over time is a question addressed in Section 6.

4.2. Tier 2: Hybrid Collaborative

Tier 2 is the most conceptually demanding tier in the HDM. Where Tier 1 assigns leadership to AI and Tier 3 assigns leadership to the human, Tier 2 requires both. At Tier 2, collaboration does not mean that both human and AI are involved. It means that the decision cannot be resolved without both. This is what the HDM means by collaboration at Tier 2: not joint presence but necessary integration. The decision requires both contributions, and neither alone is sufficient to produce a valid outcome.

The decisions that fall within Tier 2 share a specific structural condition: data without context is incomplete, and context without data is insufficient. AI identifies patterns across student performance, engagement, and progress at a scale that lies beyond the capacity of individual human observation. The educator holds contextual and relational knowledge about each student that cannot be inferred from data alone. Without AI input, the educator is deciding without the full information the decision requires. Without human input, the AI output cannot stand as a decision at all, because it lacks the interpretive and ethical authority that makes a decision valid in an educational context. These are not equivalent absences. One leaves the decision incomplete. The other leaves it without authority. The second absence is a structural claim, not a technological one: within the HDM, a decision whose validity requires human interpretive authority cannot be valid without it, regardless of how advanced AI becomes. Personalised learning, adaptive content, formative assessment interpreted across time, and student progress all share this condition. This configuration corresponds to the human-in-the-loop arrangement identified by IMDA and PDPC (2020), applied to the category of decisions where both data and contextual knowledge are constitutive of the outcome.

Authority at Tier 2 is pre-assigned to the human, as it is across all three tiers of the HDM. What the collaboration produces is not authority but the information base on which that pre-assigned authority can be fully and responsibly exercised. The human cannot exercise that authority well without AI input. The AI cannot hold that authority at all. Final judgement is exercised at the point where data and context must be integrated into a specific decision. Before that moment, both contributions are active. At that moment, the human decides. The collaboration is therefore structured and asymmetric: both contributions are necessary in process, but authority is unambiguously human at the point of decision. Human authority is final at Tier 2 not only because the HDM assigns it there, but because the decisions at this tier carry moral weight that requires a locus of accountability that only a human professional can provide.

What separates Tier 2 from Tier 1 is not the degree of human involvement but its nature. At Tier 1 the human oversees. At Tier 2 the human integrates. AI contributes patterns that are not visible at the level of individual observation. The educator contributes contextual understanding that cannot be inferred from data alone. The decision emerges from the integration of both and cannot be reduced to either. AI identifies a pattern across a student's formative assessment results. The educator interprets what that pattern means for this student, given what is known about how that student thinks, engages, and responds. The AI produced the pattern. The educator produced the decision. Remove the pattern and the decision lacks its evidential basis. Remove the educator and the pattern has no one to interpret it and no authority to act on it. Every educator who has ever looked at a spreadsheet of student scores and thought: but that is not the whole story, already understands what Tier 2 is for.

Two collapse conditions define the boundaries of this tier and serve as its diagnostic tests. If AI dominates and the human only reviews the output, the decision has migrated to Tier 1. If the human proceeds without AI input, the decision has migrated to Tier 3. Neither migration is a failure of the model. Both are visible because the model has defined what Tier 2 requires. The HDM makes migration identifiable. What separates Tier 2 from Tier 3 is not the weight of the decision but the role of AI input within it. At Tier 2 the decision cannot be formed without AI. At Tier 3 it can, and must, be made without AI determining its direction.

Tier 2 also addresses the cognitive dimension identified in Section 3.1. Without AI handling the data-processing burden, the cost of producing a fully informed decision would exceed what most educators can sustain across a working day. AI at Tier 2 reduces that cost whilst the structure preserves the educator's authority over what the information means. This is the augmentation principle from Section 3.2 operating as informed judgement: the educator decides, and the AI ensures the decision is made with the fullest possible evidential foundation.

4.3. Tier 3: Human-Led with AI Support

If Tier 1 is defined by efficiency and Tier 2 by integration, Tier 3 is defined by responsibility. Responsibility, at Tier 3, refers to the irreversible, ethically consequential, and relationally specific nature of the decisions this tier contains. These three characteristics are the forms that high consequence and deep contextual knowledge take at their most demanding expression. Curriculum design, summative assessment, ethical and relational judgement, and academic integrity belong here not because AI is incapable of contributing to these areas, but because the validity of these decisions depends on human authority. That dependency is not a policy preference. It is a structural consequence of what these decisions are. This configuration corresponds to the human-led arrangement identified across the theoretical foundations of Section 3, where the nature of the decision makes human primacy not a structural preference but a structural requirement.

The distinction between Tier 2 and Tier 3 must be stated directly. At Tier 2, AI contributes to forming the decision: the human cannot decide well without the AI's contribution, and the decision emerges from their integration. At Tier 3, AI contributes to informing the human who decides: the AI provides input, but the decision belongs to the human regardless of what that input contains. At Tier 2, AI input is constitutive. At Tier 3, it is informational. Remove AI from a Tier 2 decision and the decision is incomplete. Remove AI from a Tier 3 decision and the decision is still valid, because its validity derives from human judgement, not from data. AI improving a Tier 3 decision does not make the input constitutive. Augmentation improves the judgement. Integration requires it. The first is Tier 3. The second is Tier 2.

AI is present at Tier 3. It is not excluded. What is bounded is its role. The HDM does not claim that AI information has no effect on the educator's thinking at Tier 3. Information always enters a frame of judgement and affects it to some degree. The claim is more precise: influence without authority is not a structural problem for the HDM, because the model does not require AI to have no effect. It requires that the effect itself does not constitute authority. An educator influenced by AI-provided data has not had their authority compromised. The structure determines the locus of authority, not the source of influence. At Tier 3, the authority to determine how AI-provided information is weighted, contextualised, and acted upon rests entirely and unconditionally with the human. AI provides the input. The educator owns the interpretation.

The boundary at Tier 3 is enforced structurally, not personally. Because the decision has been identified as Tier 3 in advance, the structure signals that AI output is informational only. The educator does not need to resist AI influence in the moment. The frame has already been set by the structure. This connects directly to the meta-decision removal established in Section 3.1: the educator does not need to ask whether this is their decision to make. Tier 3 resolves the most consequential form of drift: the uncritical acceptance of AI outputs at the moment of highest stakes.

The academic integrity example deserves specific attention because it is the decision type most likely to generate a reviewer challenge. AI tools that detect patterns associated with academic integrity concerns are performing a Tier 2 function. They are contributing data to the information base. The decision about whether a breach has occurred, what it means for this student, and what consequence is appropriate is a Tier 3 decision. The two steps happen in sequence. The first is Tier 2. The second is Tier 3. In practice, institutions often treat them as a single process. The HDM's value is precisely that it separates them. That separation is not a theoretical refinement. It is a structural protection for the student.

Tier 3 expresses the augmentation principle from Section 3.2 at its most essential. The educator is freed to do what only the educator can do: exercise the irreversible, ethically grounded, relationally informed judgement that gives these decisions their validity. The base of the pyramid is defined by volume. The middle is defined by integration. The apex is defined by responsibility. The HDM does not ask institutions to choose between efficiency and human judgement. It asks them to recognise that some decisions derive their legitimacy entirely from the fact that a human made them, and to protect that legitimacy by design.

The three tiers collectively constitute a complete structure for the distribution of decision authority across AI-integrated higher education practice. They cover the full decision environment, because the two criteria that govern tier assignment apply to every decision an AI-integrated educational practice contains, and the three tiers are the complete and non-redundant set of configurations those criteria yield.

The HDM's scope within this paper is classroom-level decision-making in AI-integrated teaching practice. Decisions at the programme level, departmental level, or institutional governance level involve different actors, different accountability structures, and different consequence profiles. The HDM does not extend to those levels here. That boundary is a deliberate scope decision that future research may extend.

Educational workflows frequently contain decisions across multiple tiers within a single process. A student progress review may begin as a Tier 2 data interpretation and end as a Tier 3 welfare judgement. The HDM assigns authority at the level of each individual decision within the process, not at the level of the process as a whole.

Two forms of movement between tiers must be distinguished. Contextual reassignment is expected and principled: the same category of decision may be assigned to a higher tier in one institutional context than another because the consequence dimension has changed. Practical drift is problematic: a decision correctly assigned to Tier 2 begins to be treated as Tier 1 because AI outputs are accepted without genuine integration. The HDM is designed to prevent the second form of movement by removing the meta-decision burden.

The augmentation principle operates specifically and differently at each tier. At Tier 1 it is efficiency with oversight. At Tier 2 it is informed judgement. At Tier 3 it is freed authority. These three expressions show that augmentation is not applied once to the model as a whole. It operates at each tier according to the nature of the decisions that tier governs.

The HDM is not one possible response to the conditions Section 3 identified. It is the response those conditions require: the structure that removes the meta-decision burden bounded rationality produces, operationalises the augmentation principle across the full range of educational decisions, and derives its three tiers from the only stable configurations that the interaction of consequence and human involvement yields. Given the premises established in Sections 2 and 3, the HDM is not a useful addition to the field. It is what the field's own foundations make necessary.

5. Empirical Validation

Sections 3 and 4 established the HDM through theoretical derivation and structural presentation. This section asks a different question: does the model reflect what educators were already doing before it existed? The study that provides the answer is Chin (2025), a qualitative investigation conducted through semi-structured interviews with seven experienced educators in a Southeast Asian higher education context. The study was conducted in the absence of formal institutional guidance on AI integration, which is precisely the condition the HDM is designed to address. Nine themes emerged across three categories: pedagogical implications, operational considerations, and ethical and professional dimensions, through a thematic analysis following Braun and Clarke's (2006) six-step inductive procedure. The educators who generated those themes had no knowledge of the HDM. They were describing their own practice, drawing their own boundaries, and making their own distinctions about which decisions AI should lead, which should remain theirs, and which required both.

What the mapping in this section demonstrates is not that the HDM fits the data. It is that the data reflects the HDM: the three tiers were already present in educator practice before the model existed to name them. That is not a coincidence of design. It is the strongest form of confirmation available. Figure 3 presents the full mapping before the subsections examine each tier in turn.

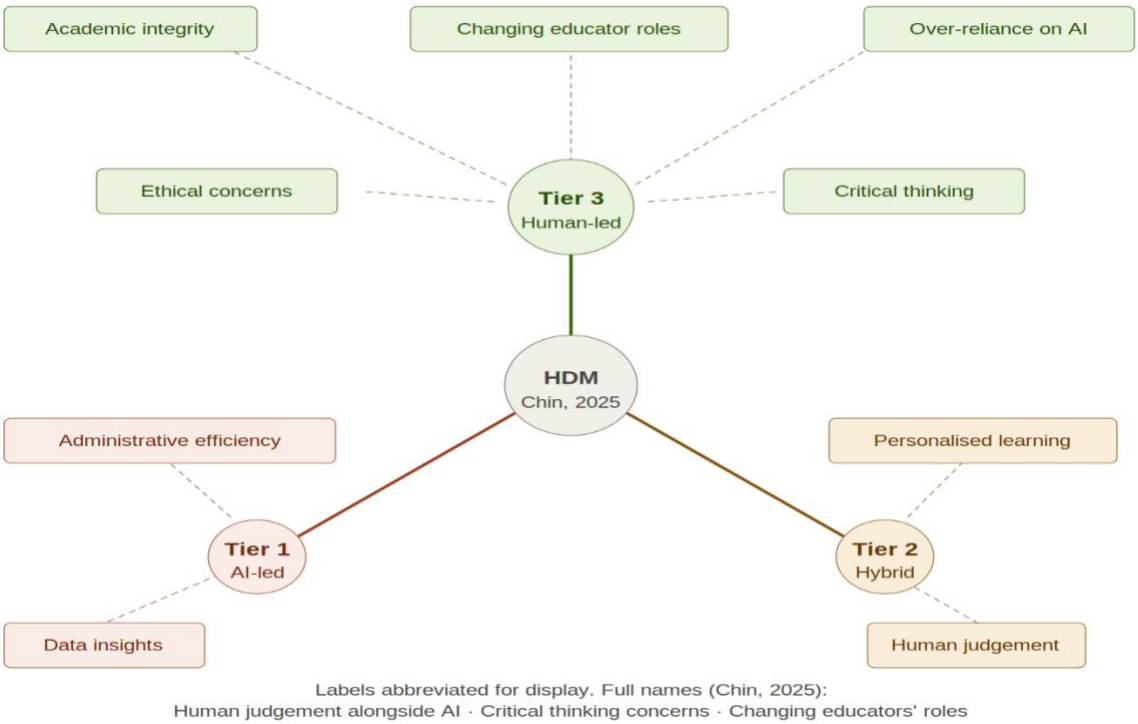

Figure 3: Hub-and-spoke diagram mapping nine empirical themes from Chin (2025) to the three tiers of the Hybrid Decision-Making Model.

Coral indicates Tier 1, amber Tier 2, and green Tier 3, with two themes mapped to Tier 1, two to Tier 2, and five to Tier 3. Abbreviated labels include Human judgement alongside AI, Critical thinking concerns, and Changing educators' roles. The mapping shows that themes from all three original categories of the study align with the HDM's tiers, indicating that the model's structure reflects patterns of practice across the dataset.

5.1. Tier 1 Validated: Administrative Efficiency and Data Insights

The two themes that map to Tier 1 are administrative efficiency and data insights, both drawn from the operational considerations category of Chin (2025). Their placement at Tier 1 rests on a structural correspondence: both themes describe decisions that are routine, repeatable, pattern-based, and low in consequence, which are precisely the conditions that define Tier 1 in the HDM.

The educators were not simply accepting AI outputs uncritically, nor were they redoing the work themselves. They were retaining the authority to review, question, and overrule whilst allowing AI to execute. Participant 1 described this explicitly: AI manages routine procedural elements whilst human expertise guides evaluation strategies. That is Tier 1 in practice. The task is delegated. The authority is not.

This pattern confirms what Tier 1 formalises. Educators had already separated execution from authority in routine decisions before any formal structure existed to require it.

Educators were consistently directing their cognitive attention towards the decisions that required it most, whilst allowing AI to handle the high-volume, pattern-based tasks that did not. That is bounded rationality working in the educator's favour: not as a source of drift, but as a natural allocation of cognitive resource that the HDM now makes structurally explicit and institutionally replicable.

5.2. Tier 2 Validated: Personalised Learning and Human Judgement Alongside AI

Two themes from Chin (2025) correspond to Tier 2: personalised learning and human judgement alongside AI. The governing claim is clear. The educators were not simply using AI to support their decisions. They were relying on AI to extend what they could know, whilst retaining full responsibility for what that knowledge meant for each student. AI revealed the pattern. The teacher decided what that pattern meant for the student. That is Tier 2, and it was already present in practice before the model existed to name it.

Personalised learning was not a general preference for individualised instruction. It was a specific practice in which AI-generated data about student performance and progress was combined with the educator's knowledge of each student as a person. AI sees the pattern across the data. The educator knows what that pattern means for this particular student. Remove the pattern and the educator is working without the information needed to decide well. Remove the educator and the pattern has no one to interpret it and no authority to act on it. This is why personalised learning belongs at Tier 2. It was placed there because the structural condition matches.

Human judgement alongside AI makes the same point through the educators' own words. What the interviews described was a practice in which AI extended what educators could know, whilst they retained responsibility for what those insights meant for each student. Participant 3 described how shifting routine tasks to AI created space for teaching decisions that depended heavily on expert knowledge and an understanding of how each student was progressing. That is not AI supporting a decision the educator could have made without it. That is AI making a form of decision possible that would otherwise be incomplete. The collaboration was not optional. It was constitutive.

A teacher who has sat with a data report at the end of a long week, trying to work out which student needs what attention next, already knows this. The data narrows the field. The knowledge of the student determines the response. Neither alone is enough.

They arrived at this arrangement through their own professional judgement about what good practice required. The HDM does not prescribe a new arrangement. It names and formalises a mode of practice that was already operating consistently in the data.

5.3. Tier 3 Validated: Academic Integrity, Ethical Concerns, Critical Thinking Concerns, Over-Reliance on AI, and Changing Educators' Roles

Five themes from Chin (2025) correspond to Tier 3, and what unites them is not their content but their consequence. Every one of them involves a moment where the educator cannot hand the decision to AI, because the decision is about a person and the educator is the one who will be held responsible for it.

The five themes are academic integrity, ethical concerns, critical thinking concerns, over-reliance on AI, and changing educators' roles. Each is a judgement about a different dimension of practice: fairness and consequence, right and wrong, student development, dependency and autonomy, and professional responsibility. But the structural condition is the same across all five. The decision cannot be reduced to data, its consequence cannot be reversed, and its validity depends on the fact that a human made it. The AI may provide a similarity score. The decision about what that score means for this student, whether a breach has occurred, what consequence is appropriate, and what the student needs, is not produced by the score. It is produced by the educator's judgement about a person, a situation, and what fairness requires. These are the conditions that Section 3.3 placed at Tier 3: irreversible consequences, ethical weight, and relational specificity that exceeds what data can capture.

The interviews in Chin (2025) show that educators were already acting as though these decisions were theirs to make. They questioned AI outputs where the consequence for the student was significant. They retained personal responsibility for judgements about academic integrity rather than treating detection tools as definitive. They described protecting students' capacity for independent thinking as a professional obligation that AI efficiency could not replace. Nobody told them to draw this boundary. They drew it because the nature of the decisions required it.

The consistency with which these five themes cluster around decisions of consequence is itself significant. The educators had no formal structure to follow. Yet they were consistently distinguishing between what AI can indicate and what only a human can decide. The boundary between AI input and human authority was already being drawn in practice before the HDM existed to name it. The earlier subsections showed that educators had already separated execution from authority and integrated data with contextual judgement. This subsection shows that they already knew which decisions were theirs alone to make. The model does not impose that boundary. It reveals the one educators were already defending.

The educators in Chin (2025) were not applying the HDM. They were responding to the structural conditions of their own practice: the reversibility of some decisions, the constitutive role of data in others, and the irreversible ethical weight of a third category. Those conditions organised their behaviour before the model existed to name them. The HDM does not describe educator behaviour. It reveals the structure that was already organising it. If a reviewer asks whether the educators knew they were operating within these tiers, the answer is no. But their decisions show that they were.

6. Reflections on Institutional Practice

Sections 3 and 4 established the HDM as a necessary structure. What remains is a different question: not whether the structure is valid, but what must be true for it to remain so. Section 3.1 noted that structural safeguards operate at the level of individual decisions, and that whether institutions maintain the conditions for those safeguards to hold over time is a question the model identifies but does not resolve within Section 4. This section addresses that question by naming the conditions the model's own logic requires.

6.1 Sustainability and the Institutional Condition

The HDM resolves drift at the level of individual decisions by removing the meta-decision burden. That resolution holds because the tier assignment is already in place before the decision arises. The same logic applies at the institutional level, and the consequences of its absence are more serious because they operate at scale.

Institutional drift accumulates through repeated decision cycles in which authority is left unclear, boundaries remain informal, and responsibility stays implicit. What accumulates is a quiet realignment of where educational judgement is exercised, away from deliberate human authority and towards whatever the AI has already produced. Decisions that belong at Tier 2 begin cycling as Tier 1. Decisions that belong at Tier 3 begin cycling as Tier 2. The structure has not broken. It has drifted.

The institutional equivalent of removing the meta-decision is the formal assignment of decision authority as part of governance, established before practice begins. Training programmes and policy statements address what educators know about the problem. They do not change the conditions that produce it. For the HDM's structural logic to hold, three conditions must be true. Decision authority must be formally assigned to tiers before educators encounter those decisions. Those assignments must be visible at the point of practice. And they must be reviewed as AI capability and context evolve. Without these conditions, the responsibility for determining decision authority returns to the individual educator, and the conditions for drift are reinstated at scale. What the HDM does for the individual, the institution must do for the system.

6.2 Responsibility and the Maintenance of Structure

The HDM transfers the responsibility for maintaining decision authority assignments to the institution. That transfer does not occur automatically. It requires institutional responsibility for monitoring whether the structure is being followed, identifying when practice has drifted, and reviewing assignments when conditions change. This is not supervision of individual decisions. It is maintenance of the structure that governs them, a governance function that operates above the level of individual practice.

A structure that is not actively maintained will not hold. Interpretation shifts. Practice changes. AI capability evolves. Without a clear structure for determining whose responsibility it is to maintain the assignment, that determination defaults to those closest to the decision at the moment it arises.

Singapore's Model AI Governance Framework (IMDA and PDPC, 2020) and Brunei Darussalam's national guide on AI governance and ethics (AITI, 2025) both treat the determination of human involvement in AI-augmented decision-making as a core organisational responsibility. Research examining how AI governance frameworks translate into policy practice confirms that transparency and accountability in decision-making remain central concerns across national contexts (Yar et al., 2024). The HDM specifies what that involvement looks like at the level of classroom practice. What those frameworks call for at the organisational level, the HDM provides at the practice level, requiring in return that institutions treat tier maintenance as a governance responsibility rather than an individual obligation.

Without a clear locus of responsibility for maintaining the structure, the assignment of decision authority gradually becomes interpretive again. The HDM removes uncertainty from decisions. Institutions must remove uncertainty from the structure that governs them over time.

6.3 Stability, Review, and the Protection of Principles

The HDM's criteria for tier assignment are stable. A decision belongs at Tier 3 not because current AI cannot handle it, but because its irreversible consequence and relational specificity make human authority structurally necessary. That structural necessity is independent of AI capability. The principles remain constant. What changes is how they apply.

As AI capability evolves and institutional contexts shift, decisions may require reclassification. That reclassification is a feature of principled application. What is a problem is when reclassification occurs informally, without being tested against the two criteria. That is not adaptation. It is unexamined reassignment of authority, and it reproduces the conditions for drift.

The HDM therefore requires deliberate review as a structural condition of its own validity. Principled reassignment is deliberate and tested. Informal reassignment is drift by another name.

Change is inevitable. Unexamined change is drift. The HDM remains valid not because it adapts to change, but because it requires its application to be deliberately reviewed so that its principles continue to govern under changing conditions. Structure must exist. Responsibility must maintain it. Review must protect it from silent change.

Section 6 has established three conditions the HDM's own logic requires. The structure of decision authority must be assigned before practice begins. That assignment must be maintained as an institutional responsibility. And its application must be deliberately reviewed as conditions change. These are not recommendations. They are the conditions without which the model's structural logic does not survive contact with institutional reality. The HDM does the work of clarifying whose decision it is. What keeps that clarity intact over time is a different order of work entirely, and it belongs to the institution.

7. Discussion

This paper argues that the central problem in AI-integrated education is not technological but structural: not what AI can do, but who decides. When AI contributes to educational decisions, authority over those decisions shifts, and that shift has largely occurred without being named or governed. Epistemic authority, as used here, refers to the recognised right to determine what counts as valid judgement in an educational setting: whose assessment stands, whose interpretation holds, and whose decision defines the outcome for the student. What the field has treated as a question of capability and risk, this paper reframes as a question of authority.

Three bodies of scholarship established that human and AI must collaborate deliberately, that the teacher must remain central, and that task allocation must be designed rather than assumed. Each stopped at the same point: they specified how work should be shared but did not specify who holds authority when a judgement must be made. The HDM begins where that stops. It provides what the existing literature established as a necessary condition but did not supply: a structure for assigning decision authority explicitly. In doing so, it completes a line of work the field had already begun, even if the field had not yet named what it was building towards.

This contribution is not only theoretical. By naming decision authority as the structural question at the centre of AI integration, it connects classroom practice directly to the governance conversations in which accountability depends on clearly defined human roles. Independent work confirms that this connection is timely: a scoping review of 32 empirical studies found that teachers were already being asked to determine what AI can and cannot do in assessment, without a structure to support that determination (Xia et al., 2024). The HDM provides that structure. For practitioners, it reframes AI use not as a matter of preference or efficiency, but as a matter of structured judgement and responsibility. The field can now move from recognising that AI affects decisions to specifying how those decisions should be governed. It can also evaluate AI use not only in terms of effectiveness or ethics, but in terms of whether decision authority is appropriately assigned.

The HDM specifies a conceptual structure for decision authority but does not prescribe mechanisms for institutional implementation. It is developed within a higher education context and may require adaptation when applied to other educational levels or sectors. Its empirical grounding in Chin (2025) confirms alignment in practice but does not establish generalisability across institutions or national settings. The framework assumes evolving AI capability but does not empirically test how technological change alters tier assignment over time. These boundaries do not diminish the contribution. They define what it sets out to do, and what it leaves, deliberately, for the work that follows.

This paper has not added another framework to a field that already has many. It has made visible the structure that was missing from all of them.

8. Future Research

The HDM makes decision authority visible. The most immediate research priority is the development of an instrument through which decision authority can be observed and measured in AI-integrated educational practice: one that allows researchers to identify which tier a decision belongs to, whether the correct authority is being exercised, and where drift is occurring without being named. Without that instrument, the questions that follow cannot be pursued empirically. What becomes possible to study now that decision authority is visible is the question this section addresses.

The first set of directions faces outward. How do institutions translate the HDM's tier assignments into governance practice? Who holds responsibility for maintaining them, and how do national frameworks such as AITI (2025) and IMDA and PDPC (2020) connect to classroom-level decision authority? These questions concern implementation and sustainability, the conditions Section 6 identified but did not resolve. They also extend to cross-national validation: whether educators in other higher education systems, working without formal guidance, draw the same structural distinctions the HDM formalises remains to be tested. And as AI capability evolves, how do institutions review tier assignments deliberately rather than allowing informal drift to substitute for principled reassignment? These are the questions closest to practice, and they are the ones the field is most immediately positioned to pursue.

The second set of directions faces inward. The HDM's three-tier structure gives researchers an instrument for observing how decision authority is actually distributed in AI-integrated classrooms, where it resides, where it has drifted, and when AI input moves from informational to constitutive without that movement being named. That observational capacity is new. Equally important is the question of whether educators feel safe exercising the authority the structure assigns. The HDM specifies that Tier 3 authority belongs to the educator unconditionally. But holding authority and feeling safe to use it are not the same condition. Whether educators feel safe challenging an AI-generated output, particularly in institutional environments where that output carries implicit endorsement, is a question the HDM makes possible to ask but does not resolve. The relationship between structural authority assignment and the psychological conditions under which that authority is exercised, explored in the context of educational leadership by Edmondson (1999) and Chin (2026), represents a necessary and largely unexamined direction for future empirical work. A structure can assign authority. It cannot assign the willingness to bear its weight. An educator who holds Tier 3 authority but does not feel safe challenging an AI-generated output has not lost the authority. The authority simply has not been used. Whether that happens, and why, is a question the HDM makes possible to ask and necessary to answer.

Funding: This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Conflicts of Interest: The author declares no conflicts of interest with respect to the research, authorship, and/or publication of this article.

Declaration of Generative AI and AI-assisted Technologies: This study has not used any generative AI tools or technologies in the preparation of this manuscript.

References

Akata, Z., Balliet, D., de Rijke, M., Dignum, F., Dignum, V., Eiben, G., Fokkens, A., Grossi, D., Hindriks, K., Hoos, H., Hung, H., Jonker, C., Monz, C., Neerincx, M., Oliehoek, F., Prakken, H., Schlobach, S., van der Gaag, L., van Harmelen, F., van Hoof, H., van Riemsdijk, B., van Wynsberghe, A., Verbrugge, R., Verheij, B., Vossen, P., & Welling, M. (2020). A research agenda for hybrid intelligence: Augmenting human intellect with collaborative, adaptive, responsible, and explainable artificial intelligence. *Computer*, 53(8), 18–28. https://doi.org/10.1109/MC.2020.2996587

Authority for Info-communications Technology Industry of Brunei Darussalam. (2025). *Guide on artificial intelligence (AI) governance and ethics for Brunei Darussalam*. AITI. https://www.aiti.gov.bn/guide-books/guide-on-ai-governance-and-ethics-for-brunei-darussalam/

Baker, R. S. (2016). Stupid tutoring systems, intelligent humans. *International Journal of Artificial Intelligence in Education*, 26(2), 600–614. https://doi.org/10.1007/s40593-016-0105-0

Borchers, C., Viberg, O., & Kizilcec, R. F. (2026). Who decides in AI-mediated learning? The Agency Allocation Framework. In Proceedings of the Thirteenth ACM Conference on Learning @ Scale (L@S '26). ACM. https://doi.org/10.1145/3774398.3811627

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Bredeweg, B., & Kragten, M. (2022). Requirements and challenges for hybrid intelligence: A case-study in education. *Frontiers in Artificial Intelligence*, 5, 891630. https://doi.org/10.3389/frai.2022.891630

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, 38. https://doi.org/10.1186/s41239-023-00408-3

Chin, P. P. L. (2025). AI in Higher Education: A Case Study Examining Decision-Making Shifts and Personalised Learning Through Educators' Perspectives. Education Quarterly Reviews, 8(2), 23–39. https://doi.org/10.31014/aior.1993.08.02.576

Chin, P. P. L. (2026). The Pause Before the Answer: Psychological Safety as the Missing Condition Between Visionary Leadership and Genuine Teacher Voice. Education Quarterly Reviews, 9(2), 1–10. https://doi.org/10.31014/aior.1993.09.02.712

Chou, C. Y., Huang, B. H., & Lin, C. J. (2011). Complementary machine intelligence and human intelligence in virtual teaching assistant for tutoring program tracing. *Computers and Education*, 57(4), 2303–2312. https://doi.org/10.1016/j.compedu.2011.06.005

Edmondson, A. C. (1999). Psychological safety and learning behavior in work teams. *Administrative Science Quarterly*, 44(2), 350–383. https://doi.org/10.2307/2666999

Engelbart, D. C. (1962). *Augmenting human intellect: A conceptual framework*. Stanford Research Institute. https://www.dougengelbart.org/pubs/augment-3906.html

Floridi, L., & Cowls, J. (2019). A unified framework of five principles for AI in society. *Harvard Data Science Review*, 1(1). https://doi.org/10.1162/99608f92.8cd550d1

Holstein, K., & Aleven, V. (2022). Designing for human–AI complementarity in K–12 education. *AI Magazine*, 43, 239–248. https://doi.org/10.1002/aaai.12058

Holstein, K., Aleven, V., & Rummel, N. (2020). A conceptual framework for human-AI hybrid adaptivity in education. In I. I. Bittencourt, M. Cukurova, K. Muldner, R. Luckin, & E. Millán (Eds.), *Artificial intelligence in education*, 240–254. Springer. https://doi.org/10.1007/978-3-030-52237-7_20

Holstein, K., McLaren, B. M., & Aleven, V. (2019). Co-designing a real-time classroom orchestration tool to support teacher-AI complementarity. *Journal of Learning Analytics*, 6(2), 27–52. https://doi.org/10.18608/jla.2019.62.3

Infocomm Media Development Authority and Personal Data Protection Commission. (2020). Model AI governance framework (2nd ed.). https://www.imda.gov.sg/resources/press-releases-factsheets-and-speeches/factsheets/2020/model-ai-governance-framework

Miao, F., & Holmes, W. (2021). *Artificial intelligence and education: Guidance for policymakers.* UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000376709

Paiva, R., & Bittencourt, I. I. (2020). Helping teachers help their students: A human-AI hybrid approach. In I. I. Bittencourt, M. Cukurova, K. Muldner, R. Luckin, & E. Millán (Eds.), *Artificial intelligence in education*, 448–459. Springer. https://doi.org/10.1007/978-3-030-52237-7_36

Popenici, S. A. D., & Kerr, S. (2017). Exploring the impact of artificial intelligence on teaching and learning in higher education. *Research and Practice in Technology Enhanced Learning*, 12, 22. https://doi.org/10.1186/s41039-017-0062-8

Schiff, D. (2022). Education for AI, not AI for education: The role of education and ethics in national AI policy strategies. *International Journal of Artificial Intelligence in Education*, 32(3), 527–563. https://doi.org/10.1007/s40593-021-00270-2

Simon, H. A. (1947). *Administrative behavior*. Macmillan.

Simon, H. A. (1990). Bounded rationality. In J. Eatwell, M. Milgate, & P. Newman (Eds.), Utility and probability, 15–18. Macmillan.

Xia, Q., Weng, X., Ouyang, F., Lin, T. J., & Chiu, T. K. F. (2024). A scoping review on how generative artificial intelligence transforms assessment in higher education. *International Journal of Educational Technology in Higher Education,* 21, 40. https://doi.org/10.1186/s41239-024-00468-z

Yar, M. A., Hamdan, M., Anshari, M., Fitriyani, N. L., & Syafrudin, M. (2024). Governing with intelligence: The impact of artificial intelligence on policy development. *Information*, 15(9), 556. https://doi.org/10.3390/info15090556