Front. Digit. Educ., 2027, 4(1): 2
https://doi.org/10.1007/s44366-026-0107-x

RESEARCH ARTICLE

Creating and Sustaining the Judgement Space: Professional Judgement and Dialogic Inquiry in AI-Supported Higher Education

Ying Ji, Neil Mercer
Hughes Hall, University of Cambridge, Cambridge CB1 2EW, UK

© The Author(s) 2026

Abstract Professional judgement, the situated, normative practice of weighing what is educationally desirable rather than merely technically effective, is central to educational professionalism. Yet international frameworks for faculty development converge on what individual faculty should know and do, without conceptualising the conditions under which professional judgement is exercised, or how AI restructures them. This paper introduces the judgement space: the pedagogical condition within which professional judgement is not only exercised but articulated, examined, and sustained through pedagogical interaction. In AI-supported higher education, where students engage with AI independently and beyond the pedagogical relationship, the judgement space becomes a three-party condition: AI-generated evaluative outputs enter the interaction alongside the judgement of teacher and student. We draw on Biesta's account of educational purpose across qualification, socialisation, and subjectification. From this account, we advance three claims. First, professional judgement is foundational to faculty development, not reducible to one competency among many. Second, AI restructures the evidence base on which professional judgement operates. It captures qualification-domain evidence more readily than evidence of socialisation and subjectification, privileging responsive practice unless the pedagogical framework positions AI to enrich rather than impoverish the judgement space. Third, the dialogic inquiry triangle, developed within the dialogic inquiry tradition and examined in school-based practitioner research, provides such a framework. Extended to AI-supported higher education, it structures the three-party interaction through which the judgement space can be created and sustained. We derive six structural principles for the teacher–student–AI interaction and a professional judgement audit for faculty development.

Keywords professional judgement, judgement space, dialogic inquiry, evaluative judgement, faculty development, AI-supported higher education

1 Introduction

International frameworks increasingly recognise that faculty development in AI-supported environments must address more than digital literacy or technical fluency with AI tools. The UNESCO's *AI competency framework for teachers* (Miao & Cukurova, 2024) defines competencies across five dimensions, specifying what teachers should know and be able to do with AI. The *OECD teaching compass: reimagining teachers as agents of curriculum change* (OECD, 2025) frames teacher development more broadly around agency, professional competencies, and well-being, calling for teachers to be empowered as co-designers of learning and adaptive experts. The *Reimagining teaching in an accelerating world: International Summit on the Teaching Profession* (OECD, 2026) goes further, asserting that AI should not replace human judgement, relationships, or responsibility in the learning process, and that professional autonomy requires systemic support to be sustained. While this aspiration is clearly articulated, the report does not specify the pedagogical conditions under which it can be realised. Existing approaches to faculty development in higher education, from pedagogical knowledge and teaching skills (Steinert et al., 2016) to more recent frameworks addressing digital competence and AI literacy (Miao & Cukurova, 2024; Redecker, 2017), similarly focus on what individual faculty members should know or be able to do.

Despite the convergence, a critical question remains: whether competencies alone are sufficient, or whether the conditions for their exercise require separate attention. In particular, professional judgement, the

Received June 1, 2026; revised August 19, 2026; accepted August 27, 2026
Ying Ji (✉)
E-mail: yj300@cam.ac.uk

situated, normative practice of weighing what is educationally desirable rather than merely technically effective, is, as Biesta (2009, 2015a) argues, constitutive of educational professionalism; it cannot be captured by any list of competencies. No existing framework provides a conceptual account of what pedagogical conditions sustain this practice or specifies how AI restructures those conditions. Scholars have begun to examine the pedagogical challenges AI presents (e.g., Bearman & Ajjawi, 2023; Bearman et al., 2024), but the conditions under which professional judgement operates in AI-supported environments have not been formalised as a construct that can be analysed, designed for, and sustained. This absence takes a distinctive form in higher education, where students engage with AI independently and beyond the pedagogical relationship, and where a third evaluative presence therefore enters the conditions for professional judgement.

This paper introduces the concept of the judgement space: the pedagogical condition within which professional judgement is exercised and sustained through pedagogical interaction. In AI-supported environments, AI generates evaluative outputs that enter the interaction alongside the judgement of teacher and student, making the judgement space a three-party condition. The paper proposes that faculty development address not only competency acquisition but also the creation of these conditions, and advances three interconnected claims. First, professional judgement is foundational to faculty development and not reducible to one competency among many. Second, AI restructures the evidence base on which professional judgement operates, in ways that can either enrich or impoverish the judgement space, depending on the pedagogical framework. Third, the dialogic inquiry triangle (hereafter the triangle), developed within the dialogic inquiry tradition (Alexander, 2020; Mercer, 2000; Rojas-Drummond et al., 2025; Wells, 1999) and examined in school-based practitioner research (Xue et al., 2025; Yang et al., 2025; Zhang et al., 2025), provides the pedagogical framework through which the judgement space can be created and sustained. We extend it here to AI-supported higher education. From this analysis, we derive six structural principles for the teacher–student–AI interaction and a professional judgement audit (PJA) for faculty development.

This paper is organised as follows. Section 2 establishes professional judgement as the core of educational professionalism; Section 3 introduces the judgement space as a formalised construct; Section 4 analyses how AI restructures the conditions for professional judgement; Section 5 presents the triangle as the pedagogical framework for creating and sustaining the judgement space; and Section 6 derives structural principles, a PJA, and a research agenda for faculty development.

2 Professional Judgement as the Practice of Educational Professionalism

Education, as Biesta (2009, 2015a) argues, functions across three inseparable domains of purpose: qualification (knowledge, skills, understanding), socialisation (becoming part of existing traditions and communities of practice), and subjectification (the emergence of the person as an autonomous subject capable of independent judgement and responsible action). These domains can both complement and conflict with each other. No educational act operates in only one domain. Even when a teacher focuses on helping students master specific knowledge, which is an apparent qualification-centred aim, they are simultaneously shaping how students relate to traditions of inquiry (socialisation) and influencing whether students experience themselves as capable, autonomous thinkers or as recipients of delivered content (subjectification). Biesta emphasises that balancing these domains is not an abstract question resolvable at the level of policy or curriculum design; it is a concrete question that returns repeatedly in practice, in relation to each individual student. This is why, on Biesta's account, education requires a particular kind of professional practice that cannot be reduced to subject knowledge, pedagogical technique, or evidence about "what works," though it draws on all of these. This practice is fundamentally normative: It requires weighing purposes and values across all three domains throughout pedagogical practice. We draw on Biesta's framework as an analytical architecture that distinguishes types of educational purpose, a distinction that becomes critical when AI restructures conditions for professional judgement differentially across these types.

Professional judgement, so understood, is always situated: It involves responding to situations that are, as Biesta (2009) puts it, "in some respect … always new and unique," requiring judgement that cannot be predetermined. This sensitivity, however, is not merely an individual attribute; it is constituted through the relational and institutional context within which the teacher works. The teacher who decides to sustain a student's struggle with a difficult concept rather than resolving it is making a judgement that privileges subjectification—the student's development as someone capable of facing difficulty and responding as themselves—over qualification, the student reaching the correct understanding efficiently. Such judgements are normative, contextual, and relational: They depend on knowing this student, this discipline, and these educational purposes.

Understanding what professional judgement addresses, however, is not sufficient; we must also understand how it operates as a practice. Allal (2013) argues that professional judgement has a dual character: It is both a cognitive act and a socially situated practice. As a cognitive act, it involves the individual teacher interpreting evidence, weighing competing purposes, and reaching evaluative conclusions. In its social dimension, it is shaped by the institutional, collegial, and disciplinary context within which teaching occurs. Wyatt-Smith et al. (2024) provide empirical support for this dual character: establishing shared understandings of quality requires both individual interpretive reasoning and collective negotiation of standards through structured professional dialogue. Both analyses develop this account in the context of assessment. We propose, however, that the dual character of professional judgement applies across all three domains of educational purpose, since the situated, normative character of judgement that Biesta describes is not inherently limited to qualification but operates wherever educational purposes are weighed.

Recognising this dual character also clarifies how professional judgement can be constrained. Biesta (2015a) identifies two interconnected forces that limit the scope for professional judgement: the culture of measurement, which narrows judgement to what can be quantified and compared, and the "what works" agenda, which reduces normative educational questions to technical ones. Biesta (2015a) observes that the professional space for teachers is currently being constructed and 'policed' in ways that often limit rather than enhance it. These forces operate across all education sectors, but take particular forms in higher education, where they converge with the student-as-consumer model (Biesta, 2015b) and the growing presence of AI-generated evaluative outputs. In this context, students' independent engagement with AI (Freeman, 2025) adds a dimension these pressures have not previously included.

The dual character of professional judgement identified by Allal (2013), combined with the forces that constrain it, has direct implications for how it is sustained. Neither individual professional development (Schön, 1992), which cultivates the cognitive dimension of judgement, nor institutional restructuring, which shapes its situated context, is sufficient on its own, because neither addresses the interaction itself: The relational encounter between teacher and student within which professional judgement is exercised, tested, and refined in real time. Biesta's account establishes why professional judgement matters and what threatens it, but does not formalise the conditions under which it operates or the space within which it occurs. We propose that sustaining professional judgement requires a pedagogical environment that supports both its cognitive and situated dimensions within the interaction, and that this environment can be specified and analysed in terms of how new forces, including AI, restructure it. The following section introduces the judgement space as a construct with specifiable properties and identifiable conditions, capable of this analytical work.

3 Judgement Space: Definition and Properties

We define the judgement space as the pedagogical condition within which professional judgement is not only exercised but articulated, examined, and sustained through pedagogical interaction. The judgement space is not a physical location, a time allocation, or a degree of professional autonomy, though each of these may shape the conditions under which it can be created and sustained; what constitutes it is the quality of the interaction itself.

Three properties, implicit in this definition, distinguish the judgement space from professional judgement understood simply as an individual act. They formalise, as properties of the condition, the features of the practice established in the preceding section. The judgement space is relational: It exists between participants in the educational interaction, not within the individual teacher alone. In AI-supported environments, it becomes a three-party condition: AI's evaluative outputs enter the interaction alongside, and are weighed against, the judgement of teacher and student. It is normative as it concerns judgement about educational purposes across all three domains, not only within the qualification domain. It is also conditional as the possibilities for its creation and sustaining are shaped by institutional, technological, and pedagogical structures that either enrich or impoverish the conditions for situated, normative judgement.

The relationship between professional judgement and the judgement space is recursive. The judgement space is the condition within which professional judgement is exercised and sustained, but it is the exercise of professional judgement that constitutes and sustains the judgement space in turn. Each cycle of this process reconstitutes the conditions for the next. In higher education, this recursive character takes on particular significance because students engage with disciplinary knowledge and AI-generated evaluative outputs both within and beyond the pedagogical relationship. Cycles of the process occur beyond the teacher's participation, and the conditions they reconstitute re-enter the pedagogical interaction from outside. More broadly, any force that restructures the conditions for professional judgement does not act at a

single moment; it enters an ongoing dynamic. When AI enters the pedagogical environment, it also enters this dynamic.

4 AI and the Conditions for Professional Judgement

AI is now positioned in a widening range of pedagogical roles in higher education, extending well beyond the evaluation of student performance (Crompton & Burke, 2023). Recent studies position AI as a Socratic interlocutor to scaffold critical thinking through questioning (Bazouche, 2025; Fakour & Imani, 2025), and systematic review evidence documents the influence of generative AI (GenAI) usage on students' self-regulated learning and critical thinking (Sardi et al., 2025). Within this widening field, our analysis centres on AI-generated evaluative outputs because it is through the evaluative dimension of AI's pedagogical roles that the conditions for professional judgement are most directly restructured. Even when AI functions as a dialogue partner rather than an assessor, its questions, prompts, and responses carry evaluative content, directing which lines of reasoning are pursued and framing what counts as an adequate response. The analysis of the judgement space therefore extends to these roles, which will be discussed in Section 5.

4.1 | Two Configurations: Judgement-Bypassing and Judgement-Enriching

How AI restructures the conditions for professional judgement depends on where AI sits in relation to that judgement, and we distinguish two configurations. In a judgement-bypassing configuration, AI-generated evaluative outputs function outside a pedagogical framework that brings them into normative deliberation, where the student receives such feedback, follows an adaptive pathway, or acts on evaluative content without conditions that subject these outputs to professional judgement. In a judgement-enriching configuration, the teacher's professional practice has created conditions within which AI outputs become material for professional judgement and for students' developing critical engagement. What distinguishes these configurations is not whether AI outputs reach students without real-time teacher mediation; in higher education, they already do: the large majority of students now use GenAI routinely and independently in their studies, including within the pedagogical setting itself (Freeman, 2025).

The distinction turns on whether the teacher's professional practice, exercised through pedagogical design, including how learning activities are framed and how AI tools are selected and configured, has created a pedagogical framework within which AI outputs are encountered as material for deliberation rather than accepted as definitive evaluation. Consider a student who uses a GenAI tool to obtain feedback on an essay. In isolation, this is a judgement-bypassing interaction. But if the teacher has designed a subsequent seminar in which students examine their AI-generated feedback against the quality criteria the discipline values, the same AI output now functions within a judgement-enriching interaction. The teacher's professional judgement operates not at the moment of AI output but in the design, framing, and facilitation of the pedagogical environment. Table 1 formalises the distinction along its two defining dimensions: the locus of professional judgement and the student's relation to evaluation.

4.2 | A Restructured Evidence Base: Epistemic Asymmetry

The judgement-enriching configuration depends on conditions that can be restructured by forces beyond the individual teacher's influence. Unlike the culture of measurement and the "what works" agenda, which limit the scope for professional judgement from outside the pedagogical interaction (Biesta, 2015a), AI enters the interaction itself, generating evaluative outputs that restructure what counts as evidence of learning, which domains of educational purpose receive sustained attention, and where the occasions for normative deliberation occur. Bearman and Ajjawi (2023) identify the core challenge: The interpretive work of engaging with AI outputs is continuous, contextual, and cannot be decomposed into discrete tasks that are delegated to AI or retained by the teacher.

This restructuring operates with a persistent epistemic asymmetry across Biesta's three domains. The qualification domain yields the type of evidence that AI can most readily capture, process, and act upon, including correct answers, error patterns, knowledge gaps, and performance metrics, since quality criteria in this

Table 1 Definitions of the judgement-bypassing and judgement-enriching configurations

Dimension	Judgement-bypassing configuration	Judgement-enriching configuration
Locus of professional judgement	Not brought to bear on the AI output within the interaction	Exercised across the design, framing, and facilitation of the pedagogical environment
Student's relation to evaluation	Acts on externally generated evaluation	Examines, contextualises, and judges the evaluation against disciplinary quality criteria

domain can be more completely defined in advance. In the socialisation and subjectification domains, AI can contribute useful data, for example through dialogic analysis tools that track interaction patterns, turn-taking, and response characteristics. However, even when such data are captured, this requires professional judgement that cannot be substituted by algorithmic processes. The meaning of a student's hesitation, risk-taking, challenge to a peer, or shift in perspective depends on normative, contextual, and relational judgements that only a teacher who knows this student, this discipline, and these educational purposes can make. We argue that this asymmetry is not a limitation of current AI systems that more sophisticated technology will resolve. It reflects the nature of socialisation and subjectification as domains, where "what counts as quality" is itself a normative question, and is inseparable from the situated judgement of the teacher.

Empirical evidence bears on this asymmetry in several ways. Fan et al. (2025), in a randomised experimental study, found that learners using GenAI improved essay scores but showed no corresponding gains in knowledge acquisition or transfer, exhibiting patterns consistent with what the authors term "metacognitive laziness," which means reliance on AI-generated guidance rather than active self-monitoring and evaluation. We interpret this as consistent with the epistemic asymmetry: Qualification-domain performance improved, while the capacity for independent self-monitoring and evaluation, the autonomous judgement that characterises subjectification, did not. Systematic review evidence points in both directions. GenAI is reported to enhance self-regulated learning largely through feedback and planning support, while concerns about over-reliance recur across the same literature (Sardi et al., 2025). Whether enrichment or impoverishment of the judgement space prevails depends, we argue, on the pedagogical framework within which AI operates. Alghamdi and Alghizzi (2025), reviewing educators' reflections across 37 empirical studies, found that AI's recognised capacity to enhance responsiveness and consistency was persistently accompanied by concerns about the weakening of teacher-student relationships: the relational dimensions through which socialisation operates.

Even when the authority of teachers' judgement is formally preserved, an evidence base systematically weighted toward qualification-domain data leaves professional judgement operating on an uneven epistemic foundation. Teachers may gravitate toward qualification concerns, simply because evidence within this domain is the richest and most readily available. The judgement space, which requires normative deliberation across all three domains, is narrowed not by removing the teacher's authority but by impoverishing the evidence on which that authority operates.

4.3 | From Responsive-only Practice to Dialogic Inquiry

The epistemic asymmetry thus creates conditions that privilege what we distinguish as responsive-only practice. AI provides what it is designed to provide: It reacts to data in real time, detects patterns, identifies issues, and generates immediate feedback with consistency and at scale. Professional judgement provides what AI, by its nature, does not: Normative commitment to educational purposes across all three of Biesta's domains, ethical accountability for the consequences of pedagogical decisions, and acceptance of the risk that the right educational decision might not be the most efficient one (Biesta, 2015a, 2015c, 2021).

Professional judgement within this judgement space must be both responsive and responsible, two qualities that operate concurrently rather than sequentially. Responsive practice attends to what is happening, reading the data, noticing the pattern, reacting to performance. Responsible practice asks what should be happening, deliberating about purposes across all three domains and judging whether the current focus is serving the educational goals that matter. We term the practice through which responsiveness is held accountable to responsibility the normative pause: a deliberate redirection of attention from reacting to evidence toward interrogating whether the evidence captures what matters across all three domains. The normative pause is not a cessation of activity; it operates both in real-time facilitation and in pedagogical design, for instance when a teacher decides whether to configure AI to provide immediate answers or to structure conditions in which students must reason through difficulty independently. In responsive-only practice, AI detects student difficulty and provides immediate resolution. In responsive-and-responsible practice, the teacher judges whether the struggle is productive (Murdoch et al., 2020), serving the student's development as an autonomous thinker, or unproductive, requiring intervention. This judgement depends on normative commitments about what education is for, commitments that differ across contexts, disciplines, and individual students. AI-supported environments, when designed without attention to these conditions, risk privileging responsiveness in ways that leave teaching technically active but normatively thin. Sustaining the normative pause and addressing the epistemic asymmetry require a specific form of pedagogical practice.

The research tradition on classroom dialogue has identified the processes through which collective understanding is constructed and refined. In the type of dialogue called exploratory talk, participants offer reasons, challenge constructively, and seek shared understanding (Mercer, 2000, 2013, 2025; Mercer &

Howe, 2012). This enables them not just to interact, but to interthink (Littleton & Mercer, 2013). We argue that such collective critical thinking is particularly necessary in AI-supported environments, rather than one pedagogical option among many, because it simultaneously addresses the epistemic asymmetry and sustains the normative pause. Professional judgement, as Section 2 established, is not only a cognitive act but a socially situated practice that is articulated and refined through interaction (Alexander, 2020; Allal, 2013). Exploratory talk and the interthinking it sustains address the epistemic asymmetry by creating conditions in which evidence across all three domains is generated, interpreted through professional judgement, and subjected to normative deliberation. When students articulate their thinking, challenge each other's reasoning, and deliberate about what counts as quality in their discipline, they make visible the processes through which socialisation and subjectification occur, creating evidence that requires professional judgement to interpret. Dialogue also sustains the normative pause, because exploratory talk structures interaction around the question of what is educationally desirable rather than what is technically indicated. Participants are required to give reasons and reach considered conclusions, rather than accepting the most immediately available responses, including AI-generated responses. Wells (1999) terms the epistemological stance underlying these processes dialogic inquiry: understanding constructed through collective investigation and deliberation. The following section presents the triangle as the pedagogical framework that structures dialogic inquiry across three dimensions, operationalising it for the creation and sustaining of the judgement space.

5 Dialogic Inquiry Triangle: Creating and Sustaining the Judgement Space

Productive dialogue requires conditions that must be deliberately structured (Alexander, 2020; Howe et al., 2019; Mercer, 2000; Mercer & Littleton, 2007). The triangle operationalises this principle across three interconnected dimensions of inquiry: INTO dialogue (developing analytical capacity), THROUGH dialogue (collaborative deliberation), and FOR dialogue (transformation orientation). The triangle was developed through the implementation of the Think-Talk Toolbox (Ji & Hennessy, 2024), a framework designed to cultivate transversal skills through inquiry and dialogue. It serves as the architecture for teacher inquiry in this implementation, which is documented in published practitioner research across contrasting school contexts (Xue et al., 2025; Yang et al., 2025; Zhang et al., 2025).

In AI-supported higher education, two contextual shifts reshape how each dimension operates. First, the inquiry takes a distributed form. While both teachers and students engage in inquiry across educational contexts, in higher education the teacher's inquiry operates primarily through pedagogical design, whereas the student's inquiry takes place through critical engagement with AI outputs, both within and beyond the conditions created by the teacher. Second, the object of inquiry changes. Participants interrogate not only dialogue processes but also the relationship between AI-generated evaluative outputs and professional judgement across all three domains. The triangle, so extended, addresses a recognised gap in higher education practice, where students' spoken participation is predominantly framed in terms of assessed products, such as oral presentations, rather than as a medium through which understanding is constructed (Heron et al., 2023) and judgement developed. In higher education seminars, productive dialogue requires reasoning, justification, and argumentation oriented toward disciplinary understanding (Heron, 2023).

The triangle's transfer to higher education cannot simply be assumed. Dialogue takes discipline-specific forms. Comparative evidence indicates that the character of educational dialogue varies across disciplinary contexts in its patterns of reasoning, justification, and challenge (Gamarra et al., 2026). This variation extends to students' engagement with AI itself. In a comparative study of ChatGPT and human tutors for critical thinking development, students' preferences varied across disciplinary contexts (Fakour & Imani, 2025). What counts as quality in exploratory talk is inseparable from the epistemic practices of the discipline concerned. We therefore propose the triangle as a framework that operates through pedagogical design within these disciplinary forms, not as a uniform interactional template. The conditions it specifies are most readily created in seminars, tutorials, laboratory discussions, and studio critiques, where sustained interaction is structurally available. In large lecture-based provision and in fully online or asynchronous programmes, the scope for exploratory talk is more constrained, and the three dimensions would need to operate through structured peer interaction, designed discussion tasks, and the framing of students' independent engagement with AI, rather than through real-time facilitation. Whether the relationships between the dimensions hold under these conditions is an empirical question the research agenda addresses.

Each dimension contributes distinctively to creating and sustaining the judgement space. Inquiry INTO dialogue develops the shared analytical capacity that professional judgement requires. It provides the

vocabulary and conceptual tools to articulate what counts as quality in a given discipline and educational context, to distinguish what AI-generated evidence captures from what it misses, and to interrogate whether the current evidence base supports professional judgement across all three domains. In AI-supported environments, INTO is where participants develop the capacity to interrogate AI outputs, asking what was measured, what quality criteria were applied, and what was missed. INTO draws on the principle that awareness of how language functions is foundational for productive talk (Mercer, 2000; Mercer & Dawes, 2014); in AI-supported higher education, this principle extends to awareness of how AI-generated evaluations function. INTO creates the conditions for professional judgement to operate on evidence across all three domains, not only on the qualification-domain data that AI makes most readily available.

Inquiry THROUGH dialogue makes professional judgement a relational practice rather than an individual act. Without THROUGH, judgement risks remaining private and unexamined. It is where interthinking (Littleton & Mercer, 2013) operates most directly. Participants construct understanding together, with insights emerging through building on, challenging, and synthesising each other's reasoning—insights that no individual working independently could produce. Evidence indicates that professional learning occurs through the discourse processes of collaborative inquiry rather than through individual reflection alone (Dudley, 2013; Ji, 2021, 2023). In AI-supported environments, THROUGH is where participants make their judgement visible to each other through exploratory talk, subjecting AI-generated evaluations, including the questions and responses of AI dialogue partners, to the collective deliberation that constitutes the relational conditions of the judgement space. The triangle adds the recognition that THROUGH requires INTO as its foundation. Without shared analytical capacity, collaborative discussion remains general rather than analytically precise. The research agenda below asks whether this foundational relationship holds in higher education, where participants bring greater disciplinary expertise and critical capacity.

Inquiry FOR dialogue orients the judgement space toward transformation and developing autonomy. Without FOR, INTO, and THROUGH may generate analysis and shared understanding but do not, on their own, change practice or develop the independent judgement toward which FOR orients both teacher and student. FOR is where the responsive/responsible distinction operates as a developmental trajectory. In AI-supported higher education, both teachers and students move along this trajectory, with the teacher developing from responsive-only toward responsive-and-responsible practice, and the student moving from dependence on external evaluation toward judging for themselves. The normative pause (Section 4) takes a specific form in FOR, where the judgement of whether to sustain productive struggle or to intervene becomes visible through dialogue about the reasoning behind pedagogical decisions. FOR ensures that the judgement space is oriented toward the educational purposes that make it worth creating, where professional agency (Priestley et al., 2015) is manifested in the capacity for autonomous professional judgement, collaborative action with colleagues, and partnership with students as participants in the judgement space.

The three dimensions are mutually constitutive: analytical capacity (INTO) requires collaborative deliberation (THROUGH) to be validated, collaborative deliberation requires analytical capacity to achieve analytical precision, and both require transformation orientation (FOR) to develop the autonomous judgement that the judgement space is designed to sustain. Figure 1 illustrates this architecture, with the overlapping regions representing the intermediate conditions that arise when two dimensions operate together and the judgement space emerging at the centre where all three converge. Creating the judgement space is not a matter of adding dialogue to an AI-supported curriculum but of structuring inquiry so that exploratory talk and interthinking are directed toward analytical capacity, collaborative deliberation, and transformation simultaneously. The triangle thus provides a pedagogical framework through which the judgement-enriching configuration can be designed and sustained across the teacher's full professional practice. The following section derives six structural principles from this analysis, specifying the conditions for the three-party interaction across Biesta's three domains, and presents the PJA as a reflective tool for faculty development.

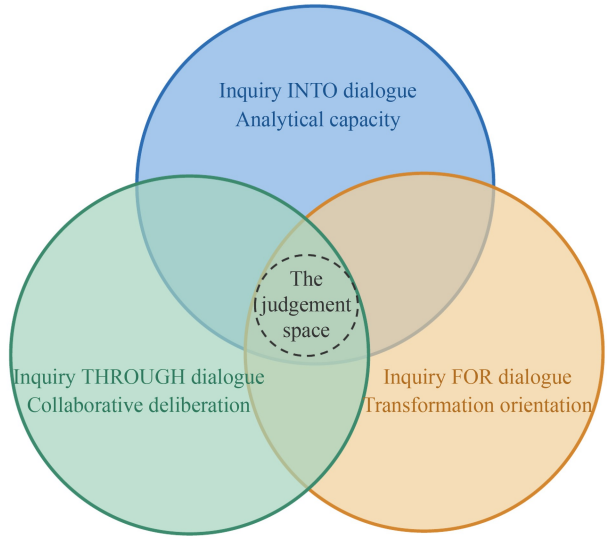


Figure 1 Dialogic inquiry triangle.

6 Implications for Faculty Development in AI-Supported Higher Education

The preceding analysis reframes how faculty development is understood. Existing approaches converge on individual knowledge and capability. We argue that this focus, while necessary, is insufficient. Faculty development in AI-supported environments must extend to creating the pedagogical conditions under which professional judgement can function across all three of Biesta's domains. These conditions also support what Tai et al. (2018) and Bearman et al. (2024) term evaluative judgement: the developing capability to make decisions about the quality of one's own and others' work, alongside critical engagement with AI outputs and deliberation about educational purposes.

6.1 | Structural Principles for the Three-Party Interaction

Mercer (2000) established that productive dialogue can be promoted by making explicit the relevant ground rules: norms that specify the conditions for productive engagement without prescribing what the dialogue concludes (Luckin et al., 2017). In AI-supported environments, an analogous set of structural principles is needed for the three-party interaction between teacher, student, and AI, ensuring the conditions for professional judgement are created and sustained.

The conditions identified in the preceding analysis can be specified as six structural principles for the three-party interaction in AI-supported higher education, organised across Biesta's three domains of educational purpose. The domains are analytically distinct but always in play simultaneously. For clarity of analysis, the principles specify conditions within each domain, addressing the two dimensions formalised in Table 1. The first concerns how professional judgement structures the interaction at the design level, positioning AI outputs within the pedagogical environment so that they enter the interaction as material for deliberation rather than as definitive evaluation. The second concerns the conditions under which students develop within the interaction toward independent judgement. Across all six principles, professional judgement is not only exercised but made visible, articulated, and opened to dialogue.

6.1.1 Qualification: How Evidence of Learning Is Interpreted and Used

Principle 1: AI provides data for professional judgement, not definitive evaluation. AI outputs are interpreted and contextualised through professional judgement before they inform pedagogical decisions. This practice takes into account both the contextual knowledge that AI does not access, including student histories and classroom dynamics, and the normative question that AI does not address—what counts as quality in this subject and this educational setting. AI outputs that function as definitive evaluation risk undermining the conditions for professional judgement. Where the pedagogical design maintains the teacher's interpretive role, AI outputs enter the judgement space as one source of evidence among several.

Principle 2: Students articulate their own learning in evaluative language that reflects their developing judgement. Students develop the capacity to describe their learning, its purposes, and its quality, rather than relying solely on externally generated scores or metrics. In AI-supported higher education, this means students can explain why certain quality criteria matter, examine what an AI-generated evaluation captured and what it missed, and articulate where their own assessment of their work differs from the AI's. This capacity for articulation and judgement is itself an educational outcome.

6.1.2 Socialisation: How Participants Engage with Perspectives, Inquiry, and Each Other

Principle 3: AI enriches the conditions for collaborative inquiry rather than replacing dialogue with individualised pathways. AI is configured to support dialogue between students and to enrich the evidence available for professional judgement about the quality of participation. However, professional judgement must attend to dimensions that AI metrics do not capture, including how students engage with different perspectives, how they negotiate meaning within the traditions of inquiry in their discipline, and whether participation reflects genuine engagement with the disciplinary community. Where AI is positioned in this way, it generates evidence that enriches the judgement space without substituting for the dialogic processes through which socialisation operates.

Principle 4: The pedagogical interaction ensures students encounter perspectives different from their own, not only AI-personalised content. Students build new understanding together through collaborative inquiry, generating insights that no individual working with AI alone could reach. In AI-supported higher education, this means the interaction is structured so that dialogue across difference, whether disciplinary, experiential, or evaluative, is a regular feature of the pedagogical environment, not an incidental outcome.

6.1.3 Subjectification: How Students Develop as Autonomous Thinkers

Principle 5: The pedagogical interaction preserves occasions for sustained productive struggle. Intellectual difficulty is not automatically resolved by AI, but is treated as an occasion for professional judgement about whether the struggle serves the student's development as an autonomous thinker. This is the normative pause in practice—the deliberate decision to sustain rather than resolve, grounded in judgement about what is educationally desirable at this moment. Where the pedagogical design creates conditions in which productive struggle is sustained, students encounter intellectual difficulty as an opportunity to develop their own capacity to face uncertainty and judge for themselves, rather than as an inefficiency to be resolved.

Principle 6: Students progressively take responsibility for their own independent judgement. Students are supported to exercise critical judgement about when AI is helpful and when it undermines their developing autonomy, moving from dependence on external evaluation, whether AI-generated or teacher-provided, toward judging for themselves. In AI-supported higher education, it means students develop the capacity not only to use AI critically but also to recognise when refraining from AI use serves their development as independent thinkers.

The recursive character of the judgement space (Section 3) means that these principles are not one-time design commitments but ongoing conditions, reconstituted through each interaction. The principles also require action beyond the individual teacher, including how AI tools are procured, how their outputs are positioned within assessment policy, how professional development equips faculty to exercise the normative pause, and how programme design sustains students' active role in the judgement space. Practitioner research from school-level implementation suggests that dialogic conditions created through the triangle support the development of evaluative judgement and critical engagement at both student and teacher levels (Xue et al., 2025; Yang et al., 2025); whether this relationship holds in AI-supported higher education, where the judgement space operates within a three-party dynamic, is a question the research agenda takes up.

6.2 | Professional Judgement Audit

The structural principles provide the normative anchor for the conditions that should be designed into the pedagogical environment, while the PJA (Table 2) provides the reflective mechanism for assessing whether these conditions are being realised in practice. The PJA is organised around Biesta's three domains of educational purpose, with three parallel perspectives on the three-party interaction: what the teacher does ("In my teaching, do I..."), what students do ("In our setting, do students..."), and how AI is positioned ("In our setting, does AI..."). Where the principles address the two dimensions formalised in Table 1, the PJA makes all three parties visible, supporting collegial discussion of how judgement is distributed across the interaction and whether any domain is normatively thin. At the bottom of each domain, a responsive/responsible contrast provides a diagnostic mirror, helping faculty recognise which description more closely matches their current practice. The PJA is designed for collegial discussion rather than individual scoring, and is presented here as a reflective tool whose development and validation as a formal instrument across disciplinary and institutional contexts is part of the research agenda proposed below.

6.3 | Evidential Basis of Framework

The framework proposed in this paper is conceptual. We state explicitly which of its elements are theoretically derived and which rest on empirical evidence. The judgement space construct, its three properties, and the six structural principles are theoretically derived from Biesta's account of educational purpose and from the dialogic inquiry tradition. They are proposed here as constructs for empirical investigation, not as established findings. The epistemic asymmetry is advanced as a conceptual claim about the differential specifiability of quality criteria across the three domains; the published empirical studies cited in Section 4 (Alghamdi & Alghizzi, 2025; Fan et al., 2025) provide evidence consistent with this claim but were not designed to test it. Systematic review evidence that GenAI's effects on self-regulated learning are mixed (Sardi et al., 2025) is consistent with the framework-dependency argument advanced in Section 4 but does not test it directly. The prevalence of independent student engagement with GenAI is established by survey evidence (Freeman, 2025). The triangle has been examined in school contexts through published practitioner research arising from its implementation (Xue et al., 2025; Yang et al., 2025; Zhang et al., 2025), which documents students' developing critical engagement and teachers' collaborative inquiry across contrasting settings. The dimension-level relationships within the triangle, including the foundational dependence of THROUGH on INTO, remain part of our ongoing empirical work. The extension of the triangle to AI-supported higher education, the three-party dynamic, and the PJA are theoretically derived and remain untested; these constructs form the core of the research agenda that follows.

Table 2 Professional judgement audit

Examine your AI-supported teaching practice by considering the prompts below. For each domain, reflect on how judgement is distributed across the three parties (teacher, student, and AI) and whether your practice is predominantly responsive (reacting to data within the qualification domain) or also responsible (deliberating about purposes across all three domains). Guiding questions: Where in my practice is professional judgement being exercised, and where has it shifted to AI? Which domain receives the most attention in my teaching? Which is thinnest? What is the difference between what I intend and what actually happens?

In my teaching, do I…	In our setting, do students…	In our setting, does AI…
Qualification—*How evidence of learning is interpreted and used*		
Interpret AI-generated assessment data through professional judgement, contextualising what it captures and what it misses, including when students have already engaged with AI outputs independently	Engage critically with AI feedback rather than accepting it as definitive evaluation	Present evaluations as material for deliberation rather than as final judgements
Discuss quality criteria as normative choices rather than fixed standards, explaining why different aspects are prioritised at different moments	Describe their progress in their own evaluative language, not only through externally generated scores or metrics	Support the teacher's interpretive role rather than bypassing it with automated actions
Responsive practice: AI scores are transmitted directly; the teacher reacts to data. Responsible practice: AI scores are interpreted through professional judgement about what constitutes quality in this disciplinary and educational context.		
Socialisation—*How participants engage with perspectives, inquiry, and each other*		
Judge the quality of participation beyond what AI metrics capture, attending to how students engage with different perspectives and negotiate meaning within the traditions of inquiry in their discipline	Encounter and engage with perspectives that differ from their own, rather than receiving AI-personalised content that narrows their exposure	Support collaborative dialogue and collective inquiry rather than individualising learning in ways that reduce genuine interaction
Build on students' contributions using disciplinary knowledge to deepen collective inquiry	Build new understanding together through dialogue, challenging each other respectfully and helping each other see things in new ways	Enrich the evidence available for professional judgement about participation quality, without replacing that judgement
Responsive practice: Participation is measured by AI-tracked frequency, word count, and response time. Responsible practice: The teacher judges whether participation reflects genuine engagement with the disciplinary community and with perspectives beyond the student's own.		
Subjectification—*How students develop as autonomous thinkers*		
Sometimes sustain productive struggle rather than resolving it, judging when difficulty serves the student's development as an autonomous thinker	Articulate what their education is for beyond grades, metrics, and employment outcomes	Sustain space for educational uncertainty and genuine intellectual risk rather than optimising toward predetermined outcomes
Make purposive trade-offs visible, explaining when and why I prioritise one educational purpose over another	Exercise critical judgement about when AI is helpful and when it undermines their own developing autonomy	Function as one resource among several that the student can engage with critically, not as the authoritative evaluator
Responsive practice: AI detects student difficulty and provides immediate resolution; the student progresses efficiently. Responsible practice: The teacher judges whether the struggle is productive, weighing responsiveness to the student's immediate needs against responsibility for their development as someone capable of facing uncertainty and judging for themselves.		

Notes. Read across each row to see how judgement is distributed in the three-party interaction for a specific domain. Read down each column to see each party's overall role across domains. The responsive/responsible contrast at the bottom of each domain is a diagnostic mirror: Consider which description more closely matches your current practice. Each domain of educational purpose corresponds to Biesta (2009, 2015a).

6.4 | Future Research Agenda

Policy frameworks across education sectors are increasingly recognising the role of teachers' professional knowledge in navigating AI-supported environments (European Commission, 2026; Miao & Cukurova, 2024; OECD, 2025, 2026), but the conditions under which professional judgement operates in these environments remain empirically underspecified, even as GenAI use among students has become near-universal (Freeman, 2025). The judgement space framework generates a research agenda with three priorities, each building on the preceding one: The principles can be evaluated only once the judgement space has been empirically investigated, and the PJA can be validated only once the principles themselves have been tested. The first priority is empirical investigation of how the judgement space operates in AI-supported higher education, asking how evaluative evidence is distributed across the three domains, whether the evidence base available to faculty privileges qualification-domain data, and how the responsive/responsible dynamic functions across disciplinary contexts. Comparative case studies across contrasting disciplines, combining analysis of the AI-generated evidence available to faculty, observation of seminars and other interactional settings, and interviews that examine the reasoning behind pedagogical decisions, offer an appropriate methodology. The second priority is to establish which of the structural principles prove most consequential for professional judgement across disciplinary and institutional contexts, and whether additional conditions beyond those specified are needed. Design-based research, in which the principles inform iterative cycles of pedagogical design across formats, from seminars and laboratories to studios and large-enrolment or online courses, is

well suited to this task, with large-scale cross-institutional programmes providing conditions for comparative investigation. Such designs would also test whether two relationships observed in school contexts hold in AI-supported higher education: the foundational dependence of collaborative deliberation on shared analytical capacity, and the contribution of the three dimensions together to students' developing evaluative judgement. The third priority is the development and validation of the PJA as a reflective instrument for faculty development, asking whether collegial use of the PJA changes how faculty design and deliberate about the three-party interaction. Longitudinal studies of the PJA within faculty development programmes across disciplines and institutions would establish its utility and validity for collegial deliberation about professional practice.

6.5 | Limitations and Conclusions

Three limitations should be acknowledged. First, the framework is grounded primarily in dialogic and philosophical literature and in empirical work conducted in school contexts; it has not yet been tested against large-scale empirical data in higher education. Second, the dialogic practices on which the framework draws carry cultural and linguistic specificity, and the conditions for exploratory talk may take different forms, or face different constraints, across the linguistic and cultural traditions within which higher education operates. Third, the structural principles assume a degree of pedagogical discretion in the design of learning environments; in resource-constrained or highly standardised institutional settings, where curricula, assessment, and AI procurement are centrally determined, the scope for creating the conditions described here may be substantially narrower. These limitations mark the boundaries of the present argument and the starting points for the research agenda above.

The preceding analysis supports three claims. First, professional judgement is foundational to faculty development, not reducible to one competency among many. Second, AI restructures the evidence base on which this practice operates, in ways that can either enrich or impoverish the judgement space depending on the pedagogical framework within which AI is positioned. Finally, dialogic inquiry, structured through the triangle, provides the pedagogical framework through which the judgement space can be created and sustained. The structural principles and PJA derived from this analysis provide practical tools for faculty development oriented toward the creation of conditions for professional judgement, not the acquisition of AI competencies alone.

Acknowledgments Open access funding was provided by the University of Cambridge.

Declaration on the Use of AI During the preparation of this manuscript, the authors used Claude (Anthropic) for language editing and refinement of draft text. The authors reviewed and edited all content and take full responsibility for the final published article.

Author Contributions Ying Ji undertook conceptualisation, development of the theoretical framework, writing of the original draft, revision, and correspondence. Neil Mercer contributed to conceptualisation, critical review and revision of the manuscript. Both authors read and approved the final manuscript.

Conflicts of Interest The authors declare that they have no conflict of interest.

Data Availability Statements No datasets were generated or analysed during the current study.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution, and reproduction in any medium or format as long as appropriate credit is given to the original authors and source, a link to the Creative Commons license is provided, and the changes made are indicated. The images or other third-party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. Visit http://creativecommons.org/licenses/by/4.0/ to view a copy of this license.

References

Alexander, R. J. (2020). *A dialogic teaching companion*. Routledge.

Alghamdi, L. H., & Alghizzi, T. M. (2025). Educators' reflections on AI-automated feedback in higher education: A structured integrative review of potentials, pitfalls, and ethical dimensions. *Frontiers in Education*, *10*, 1704820.

Allal, L. (2013). Teachers' professional judgement in assessment: A cognitive act and a socially situated practice. *Assessment in Education: Principles, Policy & Practice*, *20*(1), 20–34.

Bazouche, F. (2025). The Socratic questioning method in the age of artificial intelligence (AI) in higher education. In: Bastiaens, T., ed. *Proceedings of EdMedia + Innovate Learning 2025*. Association for the Advancement of Computing in Education, 50–58.

Bearman, M., & Ajjawi, R. (2023). Learning to work with the black box: Pedagogy for a world with artificial intelligence. *British Journal of Educational Technology*, *54*(5), 1160–1173.

Bearman, M., Tai, J., Dawson, P., Boud, D., & Ajjawi, R. (2024). Developing evaluative judgement for a time of generative artificial intelligence. *Assessment & Evaluation in Higher Education*, *49*(6), 893–905.

Biesta, G. (2009). Values and ideals in teachers' professional judgement. In: Gewirtz, S., Mahony, P., Hextall, I., & Cribb, A., eds. *Changing teacher professionalism: International trends, challenges and ways forward*. Routledge, 204–213.

Biesta, G. (2015a). What is education for? On good education, teacher judgement, and educational professionalism. *European Journal of Education, 50*(1), 75–87.

Biesta, G. (2015b). *Good education in an age of measurement: Ethics, politics, democracy*. Routledge.

Biesta, G. (2015c). *The beautiful risk of education*. Routledge.

Biesta, G. (2021). *World-centred education: A view for the present*. Routledge.

Crompton, H., & Burke, D. (2023). Artificial intelligence in higher education: The state of the field. *International Journal of Educational Technology in Higher Education*, *20*(1), 22.

Dudley, P. (2013). Teacher learning in Lesson Study: What interaction-level discourse analysis revealed about how teachers utilised imagination, tacit knowledge of teaching and fresh evidence of pupils learning, to develop practice knowledge and so enhance their pupils' learning. *Teaching and Teacher Education*, *34*, 107–121.

European Commission. (2026). *Guidelines for teachers and educators on tackling disinformation and promoting digital literacy through education and training*. Publications Office of the European Union.

Fakour, H., & Imani, M. (2025). Socratic wisdom in the age of AI: A comparative study of ChatGPT and human tutors in enhancing critical thinking skills. *Frontiers in Education*, *10*, 1528603.

Fan, Y. Z., Tang, L. Z., Le, H. X., Shen, K. J., Tan, S. F., Zhao, Y. Y., Shen, Y., Li, X. Y., & Gašević, D. (2025). Beware of metacognitive laziness: Effects of generative artificial intelligence on learning motivation, processes, and performance. *British Journal of Educational Technology*, *56*(2), 489–530.

Freeman, J. (2025). *Student generative AI survey 2025*. HEPI Policy Note 61. Higher Education Policy Institute.

Gamarra, E., Heron, M., Tenenbaum, H., & Baker, L. (2026). Disciplinary differences in educational dialogue: A comparative analysis of Foundation Year classrooms. *Language and Education, 40*(4), 853–872.

Heron, M. (2023). Dialogic dimensions of seminars in higher education. In: Thwaite, A., Simpson, A., & Jones, P., eds. *Dialogic pedagogy: Discourse in contexts from pre-school to university*. Routledge, 220–234.

Heron, M., Baker, S., Gravett, K., & Irwin, E. (2023). Scoping academic oracy in higher education: Knotting together forgotten connections to equity and academic literacies. *Higher Education Research & Development*, *42*(1), 62–77.

Howe, C., Hennessy, S., Mercer, N., Vrikki, M., & Wheatley, L. (2019). Teacher–student dialogue during classroom teaching: Does it really impact on student outcomes? *Journal of the Learning Sciences, 28*(4–5), 462–512.

Ji, Y. (2021). Does teacher engagement matter? Exploring relationship between teachers' engagement in professional development and teaching practice. *International Journal of TESOL Studies*, *3*(4), 42–60.

Ji, Y. (2023). The effect of teacher engagement factors on teachers' lesson study professional development in Chinese school settings. *International Journal for Lesson and Learning Studies*, *12*(4), 273–287.

Ji, Y., & Hennessy, S. (2024). *Think-talk toolbox: Cultivating transversal skills through inquiry and dialogue*. Camtree.

Littleton, K., & Mercer, N. (2013). *Interthinking: Putting talk to work*. Routledge.

Luckin, R., Baines, E., Cukurova, M., Holmes, W., & Mann, M. (2017). *Solved! Making the case for collaborative problem-solving*. Nesta.

Mercer, N. (2000). *Words and minds: How we use language to think together*. Routledge.

Mercer, N. (2013). The social brain, language, and goal-directed collective thinking: A social conception of cognition and its implications for understanding how we think, teach, and learn. *Educational Psychologist*, *48*(3), 148–168.

Mercer, N. (2025). *Oracy: The transformative power of finding your voice*. The Bodley Head.

Mercer, N., & Dawes, L. (2014). The study of talk between teachers and students, from the 1970s until the 2010s. *Oxford Review of Education*, *40*(4), 430–445.

Mercer, N., & Howe, C. (2012). Explaining the dialogic processes of teaching and learning: The value and potential of sociocultural theory. *Learning, Culture and Social Interaction, 1*(1), 12–21.

Mercer, N., & Littleton, K. (2007). *Dialogue and the development of children's thinking: A sociocultural approach*. Routledge.

Miao, F. C., & Cukurova, M. (2024). *AI competency framework for teachers*. UNESCO.

Murdoch, D., English, A. R., Hintz, A., & Tyson, K. (2020). Feeling heard: Inclusive education, transformative learning, and productive struggle. *Educational Theory*, *70*(5), 653–679.

OECD. (2025). *OECD teaching compass: Reimagining teachers as agents of curriculum change*. OECD.

OECD. (2026). *Reimagining teaching in an accelerating world: International Summit on the Teaching Profession*. OECD.

Priestley, M., Biesta, G., & Robinson, S. (2015). *Teacher agency: An ecological approach*. Bloomsbury.

Redecker, C. (2017). *European framework for the digital competence of educators: DigCompEdu*. Publications Office of the European Union.

Rojas-Drummond, S., Rubio-Jiménez, A. L., & Trigo-Clapés, A. (2025). *Learning together: Fostering communities of dialogic inquiry*. Springer.

Sardi, J., Darmansyah, Candra, O., Yuliana, D. F., Habibullah, Yanto, D. T. P., & Eliza, F. (2025). How generative AI influences students' self-regulated learning and critical thinking skills? A systematic review. *International Journal of Engineering Pedagogy*, *15*(1), 94–108.

Schön, D. A. (1992). *The reflective practitioner: How professionals think in action*. Routledge.

Steinert, Y., Mann, K., Anderson, B., Barnett, B. M., Centeno, A., Naismith, L., Prideaux, D., Spencer, J., Tullo, E., ... Dolmans, D. (2016). A systematic review of faculty development initiatives designed to enhance teaching effectiveness: A 10-year update: BEME Guide No. 40. *Medical Teacher*, *38*(8), 769–786.

Tai, J., Ajjawi, R., Boud, D., Dawson, P., & Panadero, E. (2018). Developing evaluative judgement: Enabling students to make decisions about the quality of work. *Higher Education*, *76*(3), 467–481.

Wells, G. (1999). *Dialogic inquiry: Towards a sociocultural practice and theory of education*. Cambridge University Press.

Wyatt-Smith, C., Adie, L., & Harris, L. (2024). Supporting teacher judgement and decision-making: Using focused analysis to help teachers see students, learning, and quality in assessment data. *British Educational Research Journal*, *50*(3), 1420–1448.

Xue, C., Zhao, P., & Li, M. (2025). *Promoting the development of primary school students' sense of innovation through challenging*. Camtree.

Yang, J., Tian, K., & Liu, Y. (2025). *Action research using Think-Talk Toolbox to complement developmental skills and improve the effectiveness of group cooperative learning*. Camtree.

Zhang, D., Zhu, C., & Bo, K. (2025). *Using "thinking-dialogue" to improve the questioning ability of primary school students*. Camtree.