

**RESEARCH
AS SYSTEM
INFRASTRUCTURE:
AN INTEGRATED
FRAMEWORK FOR
FURTHER EDUCATION
AND SKILLS**

CONTENTS

INTRODUCTION	03
SECTION 1: WHY RESEARCH MATTERS IN FE AND SKILLS	06
SECTION 2: THEORETICAL FOUNDATIONS	15
SECTION 3: A FIVE-TIER TAXONOMY OF RESEARCH-RELATED PRACTICE	17
SECTION 4: THE THREE-AXIS MODEL OF RIGOUR, TRANSFERABILITY AND CODIFICATION	21
SECTION 5: ORGANISATIONAL IMPLICATIONS: LEADERSHIP, GOVERNANCE AND SYSTEM STEWARDSHIP	27
APPENDIX A THE DETAILED TAXONOMY	32
REFERENCES	41

INTRODUCTION

Recent policy signals, such as Ofsted's Areas of Research Interest (Ofsted 2026), articulate thematic questions about inclusion, transitions, professional learning and post-16 pathways where stronger evidence is needed to improve outcomes for learners and strengthen system understanding. Ofsted's articulation of research priorities should not be read as a narrow inspection agenda, but as a wider signal that improvement increasingly depends on evidence that is credible, usable and embedded in professional practice.



This report sets out an integrated framework for understanding research as system infrastructure in further education (FE) and skills. However, if research is to function as system infrastructure, it must be legible to leaders and boards as well as meaningful to practitioners. A key design principle of this report is therefore to offer a framework that can be used at multiple levels: as a high-level leadership model and as a detailed professional taxonomy. It positions research not as a discrete academic function, but as a continuum of professional knowledge activity through which the sector generates insight, tests practice, builds capability, strengthens professionalism, and mobilises learning at scale. It synthesises a FE and skills research taxonomy into a coherent, navigable organisational framework. In doing so, it makes explicit the theoretical and professional foundations underpinning the taxonomy, and demonstrates how reflective practice, inquiry, innovation, evaluation and evidence synthesis operate as connected components of a self-improving system. It is written from the perspective of ETF's role in supporting professional standards, workforce development and system learning across FE and skills.

FE and skills occupies a distinctive and strategically critical position within the UK education system. It operates at the intersection of education, labour market policy, social mobility, community regeneration and place-based economic development. It serves some of the most diverse learner cohorts in the system, including young people for whom compulsory schooling has not worked well, adults returning to learning, individuals reskilling in response to technological and economic change, and communities experiencing long-term disadvantage. This agenda is particularly urgent now. FE and skills is operating in a period of sustained reform characterised by rapid technological change, workforce fragility, evolving employer demand, and increasingly complex learner participation patterns linked to modularisation, lifelong learning and flexible provision. These reforms are unfolding alongside an accountability environment in which data is often treated as a proxy for quality, and where performance narratives can outpace professional learning. In this context, research capacity is not an optional enhancement. It is a core condition of system resilience: the mechanism through which organisations make sense of complexity, avoid repeating ineffective interventions, and build credible, cumulative improvement over time.

Despite the centrality FE and skills occupies, it has historically lacked an embedded and self-sustaining research culture comparable to that found in higher education or, increasingly, in schools. Research has often been conducted on the sector rather than in or with it, externally commissioned, episodic in nature and weakly connected to professional ownership or organisational learning (Bathmaker, 2013; Keep, 2018). As a result, valuable professional knowledge frequently remains tacit, localised and under-recognised, while policy and improvement agendas are driven by metrics rather than insight. This is not only a matter of culture or conceptual framing: it is also shaped by the structural conditions in which FE and skills operates, including constrained funding, rapid reform cycles, uneven institutional capacity, and limited protected time for inquiry.

This report advances a research taxonomy and integrated knowledge-to-impact framework that reconceptualises research in FE and skills as a continuum of professional knowledge activity rather than a narrow academic function. Drawing on theories of professional knowledge, organisational learning, democratic professionalism and systems improvement, it demonstrates how reflective practice, professional inquiry, innovation, evaluation, practitioner research, applied research and evidence synthesis form a connected ecosystem for improvement (Eraut, 1994; Polanyi, 1966; Senge, 2006).

While many of the concepts within this framework are well established within educational research traditions, the distinctive contribution of this report lies in how they are integrated and repositioned for FE and skills leadership audiences.

Specifically, the framework makes three strategic moves:



- **first**, it reframes research as a continuum of professional knowledge activity rather than a specialist academic domain.



- **second**, it introduces a practical three-axis diagnostic model through which leaders can identify imbalance patterns and blockages in organisational learning.



- **third**, it explicitly foregrounds codification and mobilisation as core mechanisms of impact, recognising that evidence has limited value unless it becomes usable professional infrastructure.

At the heart of the framework is a three-axis model of rigour, transferability and codification. This model provides a diagnostic and developmental tool through which organisations, networks and professional bodies can identify strengths, gaps and blockages in their research and improvement activity. It moves the conversation beyond simplistic hierarchies of what counts as real research towards a more mature understanding of how professional insight becomes validated, shareable and system-shaping knowledge (Nonaka and Takeuchi, 1995; Petticrew and Roberts, 2006). Taken together, the taxonomy and the three-axis model position FE and skills not as a passive recipient of external evidence, but as a knowledge-producing professional system capable of generating, validating and mobilising evidence in ways that strengthen quality, professionalism and learner outcomes.

HOW TO USE THIS REPORT

This report is designed to be practical as well as conceptual. The intention is not to academicise practice, but to make the sector's existing knowledge activity visible, credible and strategically usable. The report can be read sequentially, but it is also structured so that different audiences can engage with the sections most relevant to their role.

Section 1

sets out the strategic case for research capacity as system infrastructure and explains why current practice is often fragmented. Worked organisational examples are included to demonstrate how the taxonomy supports proportionate decision-making.

Section 2

establishes the theoretical foundations and clarifies how quantitative methods and evidence review function across the continuum.

Section 3

presents the simplified five-tier taxonomy and explains how leaders can use it to commission, support and mobilise research-related activity. The full detailed taxonomy (Levels 1–14), for practitioner and research audiences requiring greater specificity, can be found in Appendix A.

Section 4

sets out the three-axis model (rigour, transferability and codification), which is the core diagnostic tool of the framework.

Section 5

explores the organisational implications for leadership, governance and system stewardship.

Appendix A

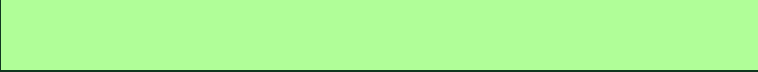
contains the full detailed taxonomy (Levels 1–14), supporting practitioner and research audiences requiring greater methodological specificity.

For accessibility and usability across different audiences, the taxonomy is presented in two complementary formats. The main body of this report sets out a simplified five-tier model designed for leadership, governance and policy audiences, intended to bring strategic clarity. Leaders should view the five tiers as a maturity map: the question is not whether an organisation is engaging with research, but whether it has the infrastructure to move learning from tacit insight into transferable and codified professional knowledge. The full detailed taxonomy (Levels 1–14), which provides greater methodological specificity and depth for practitioners and researchers, is retained in Appendix A. This two-format approach reflects the report's core principle: research culture must be both strategically intelligible and professionally authentic.

SECTION 1:

WHY RESEARCH MATTERS

IN FE AND SKILLS



The FE and skills sector face a uniquely complex improvement challenge. Providers must respond simultaneously to policy reform, economic demand, employer expectations, community need and individual learner circumstance.

They operate under intense accountability pressures while working with learners whose educational trajectories have often been fragmented or disrupted. Ofsted (2026) signals that improvement will increasingly depend on evidence that addresses inclusion, professional learning and learner pathways, emphasising the importance of embedded research practice across the FE and skills continuum. In such contexts, improvement cannot be achieved through compliance alone. While data monitoring and performance management play an important role, they are insufficient to address the relational, pedagogical and organisational complexities of FE and skills provision. What is required is organisational learning capacity: the ability of organisations and systems to inquire into their own practice, to test assumptions, to adapt intelligently and to learn collectively over time (Senge, 2006).

This framing is aligned with the arguments set out in the Education + Training special edition on why research in FE and skills matters, which positions FE and skills research as a mechanism for advancing pedagogy, professionalism and system-level improvement. Smith et al (2026) in their editorial highlights the importance of shifting from research conducted on the sector towards research conducted in and with it, strengthening ownership, relevance and impact. The framework presented here provides a practical structure for delivering that shift, supporting the sector to develop the infrastructure required for sustained inquiry, dissemination and knowledge mobilisation.

Historically, FE research has been characterised by three persistent weaknesses. First, it has often been externally driven, with research agendas set by funders, inspectors or policymakers rather than practitioners and organisations themselves. Second, it has been episodic, responding to discrete initiatives rather than embedded as an ongoing professional activity. Third, it has been disconnected from professional ownership, limiting its impact on day-to-day practice and decision-making (Keep, 2018). This has contributed to a narrative of FE and skills as research-poor, despite being rich in tacit expertise, local innovation and adaptive practice. The distinction between research conducted on the sector and research generated within it reflects deeper questions of power, legitimacy and professional identity (Sachs, 2016). When research is experienced as something done to the sector, it is unlikely to be owned, trusted or sustained.

Practitioner inquiry and practice-engaged research provide an essential counterbalance to performative accountability regimes characterised by an over-reliance on metrics and comparison (Ball, 2015). When practitioners are enabled to investigate their own practice, to generate evidence that matters locally and

to connect that evidence to wider professional knowledge, research becomes a mechanism for agency, professionalism and ethical responsibility. When infrastructure, time and capability are in place, research supports FE and skills to operate as a self-improving professional system rather than a compliance-driven delivery apparatus. It allows organisations to move from reactive problem-solving to anticipatory learning, and from isolated pockets of innovation to shared and sustainable improvement.

Throughout this report, the term research-related practice is used deliberately. It refers to the full ecosystem of activity through which FE and skills generate and mobilises professional knowledge. This includes reflective practice, professional inquiry, innovation, quality improvement, evaluation, practitioner research, applied research and academic research. These activities are related, but they are not interchangeable. They serve different purposes, require different forms of evidence, and produce different kinds of knowledge outputs. The value of the taxonomy is that it enables leaders to distinguish between them, commission proportionately, and build coherent pathways from local insight to system learning.



RESEARCH, PROFESSIONALISM AND POLICY REFORM IN FE AND SKILLS

The case for a more explicit and embedded research culture in FE and skills is not merely an abstract or academic one. It is closely bound up with current and emerging policy priorities relating to workforce reform, productivity, lifelong learning and place-based economic development. In this sense, research capacity should be understood not as an optional enhancement but as a form of system infrastructure, enabling the sector to respond intelligently and ethically to sustained reform pressures. Taken together, these reform pressures reinforce a central claim of this report: without inquiry capability and knowledge mobilisation infrastructure, FE and skills risks responding to complexity through compliance, rather than through learning.



WORKFORCE REFORM AND PROFESSIONAL CAPABILITY

Across FE and skills, concerns about workforce supply, retention and professional identity are persistent. Practitioners are increasingly expected to navigate complex roles that combine subject expertise, pedagogy, assessment, learner support and employer engagement. At the same time, opportunities for structured professional learning are often constrained by workload, funding and organisational capacity. Within this context, research-engaged practice provides a critical mechanism for strengthening professional capability without relying solely on externally delivered training. Practitioner inquiry, professional learning communities and development-oriented evaluation allow staff to deepen expertise through purposeful investigation of their own practice, anchored in real organisational challenges (Eraut, 2004; Hargreaves and Fullan, 2012). The taxonomy supports a reframing of workforce development away from episodic interventions towards sustained professional knowledge building.



LIFELONG LEARNING, MODULARISATION AND LEARNER COMPLEXITY

Policy agendas linked to lifelong learning, flexible provision and modularisation place additional demands on FE and skills providers. Learners increasingly move in and out of education, combine work and study, and engage with provision in non-linear ways. These patterns challenge traditional assumptions about curriculum coherence, progression and success. A robust research framework enables organisations to treat these challenges as sites of inquiry rather than problems to be managed reactively. Professional inquiry and evaluation can support rapid learning about learner experience, engagement and progression in new delivery models, while protecting against premature standardisation or over-reliance on blunt performance indicators. The taxonomy is therefore directly relevant to system learning in the context of lifelong learning reform.



PRODUCTIVITY, PLACE AND EMPLOYER ENGAGEMENT

FE and skills is frequently positioned as a key lever for productivity and local economic development. However, the relationship between provision, skills utilisation and economic outcomes is complex and mediated by local labour market conditions. In this context, simplistic attribution of impact risks distorting practice and undermining credibility. Research-informed approaches enable more nuanced understanding of how programmes interact with employer demand, learner aspiration and local opportunity structures. Evaluation and applied research, informed by theory and local knowledge, are essential in making credible claims about impact while avoiding over-extension. The taxonomy provides a basis for aligning ambition with proportionate evidence.



ACCOUNTABILITY, INSPECTION AND SYSTEM TRUST

Finally, the current accountability environment places significant interpretive pressure on data, inspection judgements and performance narratives. Where research literacy is weak, organisations are more vulnerable to defensive compliance behaviours and performative responses (Ball, 2015). Strengthening research capacity supports more confident, evidence-informed leadership. It enables organisations to contextualise data, articulate learning from inquiry, and engage constructively with accountability processes. In this way, research functions not only as an improvement tool but as a means of rebuilding professional trust and legitimacy within the system. Viewed through this lens, the development of a shared research taxonomy is a strategic response to policy reform across FE and skills. It supports professionalism, guards against misuse of evidence, and enables learning-led rather than compliance-led change. Taken together, these drivers suggest that research capacity is not simply a technical function, but a strategic enabler of workforce resilience, ethical practice and system legitimacy.





THE PROBLEM: FRAGMENTED CONCEPTS OF RESEARCH AND IMPROVEMENT

Across FE and skills, concepts such as innovation, evaluation, quality improvement, action research, professional inquiry and research are frequently used interchangeably. Action research is particularly significant in FE and skills because it operates as a bridge methodology across the continuum. It is simultaneously a form of professional inquiry, a mechanism of improvement, and a structured research approach.

Its distinctive value lies in the way it connects practice noticing, intervention design, implementation, reflection and evidence-building into an iterative cycle. In this sense, action research is not a discrete tier of activity. It is a methodological engine through which practice becomes professional knowledge. And while these activities share family resemblances, they are not synonymous.

Conceptual blurring generates significant practical risks. When distinctions are unclear, organisations may:

- expect inappropriate levels of methodological rigour from small-scale improvement activity
- undervalue practitioner inquiry because it does not resemble academic research
- commission evaluation that is ill-matched to purpose or stage of development
- struggle to mobilise learning beyond the immediate context in which it was generated (Cohen, Manion and Morrison, 2018).

A particular risk in FE and skills is that research can be perceived as abstract, external or disconnected from the day-to-day realities of teaching and supporting learners. This is amplified by the sector's dual professional workforce, whose expertise is shaped not only by educational practice but by occupational and industry identity. In this context, action research and practitioner inquiry provide an especially effective vehicle for engagement because they produce practice-near evidence that addresses the immediate pedagogical and operational questions practitioners face. When structured well, action research therefore functions as a safeguard against fragmented improvement: it enables evidence to be generated in forms that are meaningful, credible and usable for the workforce, while also supporting codification and dissemination beyond the originating setting.

Without a shared taxonomy, leaders lack a common language for deciding what kind of inquiry is needed, for what purpose, and to what standard. This, in turn, weakens strategic decision-making, capability development and investment. A shared conceptual framework functions as enabling infrastructure. It does not prescribe a single model of research, but rather supports proportionate, intentional and coherent approaches to inquiry across organisations and systems. In this sense, the taxonomy is not an academic artefact but a leadership and governance tool (Greany and Higham, 2018).

The following worked examples illustrate how conceptual confusion arises in real FE contexts, and how the taxonomy and three-axis model enables more proportionate, effective and professionally owned approaches to inquiry.

EXAMPLE 1

Curriculum reform is recurrent across FE and skills; as such it provides a useful lens through which to explore the practical value of the research taxonomy. Across FE and skills, curriculum change is frequently driven by a combination of qualification reform, inspection findings, employer feedback and learner outcome data. These drivers are often interpreted as requiring rapid action, leading organisations to implement new pedagogical approaches or curriculum structures at pace, frequently under the banner of innovation. In the absence of a shared conceptual framework, curriculum reform activity is often mischaracterised as research-led or evidence-based without sufficient clarity about what this means in practice. For example, an organisation may adopt a new teaching approach (such as contextualised maths, blended delivery or project-based learning) on the basis of external reports, inspection commentary or anecdotal success in another provider. The activity may then be labelled as innovation or action research, with expectations that it will demonstrate measurable impact within a short timeframe.

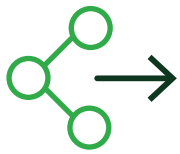
This creates several predictable problems. First, staff may be expected to generate levels of evidence that are disproportionate to the nature of the activity, leading to frustration and disengagement and circumspection about the evidence base. Second, early-stage practice development may be judged prematurely against summative outcome metrics, rather than formative indicators of feasibility or acceptability. Third, the absence of explicit evaluation design limits the organisation's ability to learn systematically from the change, even where local improvements are visible.

From a commissioning perspective, government and leaders may request research evidence of impact when what is realistically required at this stage is structured professional inquiry or formative evaluation. This mismatch between purpose and method undermines both credibility and confidence in research activity across the organisation and serves to undermine the findings and potentially the curriculum reform itself.

Using the research taxonomy allows curriculum reform to be reframed as a staged process of professional knowledge development rather than a single research event. Initial activity sits legitimately within the domains of reflective practice, professional dialogue and professional inquiry, where practitioners explore problems of practice, draw on prior knowledge, and adapt approaches to local learner cohorts (Schön, 1983; Wenger, 1998). At this stage, evidence is primarily qualitative and developmental, focusing on questions such as:

- Is the approach workable in our context?
- How are learners experiencing it?
- What adaptations are required by different teaching teams?

As confidence grows, the organisation may then move into practice development and innovation, intentionally designing curriculum models and articulating underpinning assumptions about pedagogy and learning. Here, the taxonomy supports leaders to recognise that rigour is increasing, but that claims of effectiveness should remain provisional. Only once an approach has stabilised does it become appropriate to commission evaluation or piloting, using mixed methods to test effectiveness and transferability. Quantitative analysis may now play a stronger role, but always in alignment with clearly articulated questions. At no point does the taxonomy require early-stage inquiry to masquerade as academic research; instead, it legitimises proportionate forms of evidence at each stage of development.



APPLYING THE THREE-AXIS MODEL

Mapping this work against the three-axis model provides further diagnostic clarity. Early curriculum inquiry may score highly on contextual relevance but low on codification, as knowledge remains embedded in teams rather than documented. Evaluation activity may demonstrate increasing rigour but still lack transferability if contextual conditions are not articulated. The model makes visible where leadership attention is required. For example, if innovation activity is strong but poorly codified, the response is not to increase methodological demands but to invest in documentation, peer learning and professional dialogue. Where evaluation is rigorous but disconnected from practice, attention should shift to knowledge mobilisation rather than further data collection. The final stage involves codification, where curriculum learning is translated into shared artefacts: curriculum frameworks, pedagogical guidance, CPD materials and mentoring practices. At this point, local inquiry contributes to organisational memory and, potentially, to wider system learning through networks and professional bodies (Nonaka and Takeuchi, 1995). Crucially, codification does not imply rigidity. Instead, it provides a stable reference point against which future adaptation and inquiry can occur. In this way, the taxonomy supports curriculum reform as an iterative, professionally owned process that balances responsiveness with rigour.

This worked example illustrates how the taxonomy and three-axis model enable leaders to commission inquiry proportionately, protect professional agency and strengthen the link between local improvement and system knowledge building, without imposing inappropriate academic expectations on early-stage practice development.



EXAMPLE 2

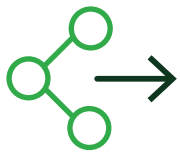
Learner retention and progression represent one of the most persistent and high-stakes challenges in FE and skills. Outcomes in this domain are closely scrutinised by boards, inspectors, funders and policymakers, and are frequently used as proxy indicators of organisational quality. At the same time, patterns of withdrawal, non-completion and irregular participation are often shaped by factors beyond the immediate control of providers, including employment volatility, caring responsibilities, health, housing insecurity and prior educational experience. In the absence of a shared framework for inquiry, challenges relating to learner retention are frequently addressed through a narrow performance-management lens. Falling attendance or completion rates are often treated as signals of weak practice requiring rapid intervention, leading to the introduction of new monitoring systems, attendance policies or learner support initiatives. These responses are commonly described as evidence-based on the grounds that they respond to data trends, even where the underlying causes of learner disengagement are poorly understood.

This dynamic creates three recurrent problems. First, correlation is mistaken for causation: changes in retention data are attributed to specific aspects of provision without adequate exploration of learner experience or contextual factors. Second, interventions are evaluated prematurely against headline outcomes, discouraging iterative learning and adaptation. Third, staff engagement with inquiry is weakened, as practitioners experience retention work as compliance activity rather than professional learning.

From a commissioning perspective, government and leaders may again default to requesting what works evidence or external exemplars, without clarifying whether the need is diagnostic understanding, practice development or impact evaluation. As with curriculum reform, this mismatch between purpose and method undermines the credibility and usefulness of improvement activity. Through the lens of the research taxonomy, learner retention can be reframed as a complex problem of practice requiring layered forms of inquiry. Initial activity rightly sits within reflective practice and professional dialogue, where teams examine patterns of learner withdrawal, triangulate quantitative data with practitioner insight, and surface assumptions about learner motivation and behaviour. At this stage, professional inquiry focuses on sense-making rather than solution generation. Key questions include:

- What are learners telling us about why they disengage?
- How do programme design, assessment practices and timetabling interact with learners' lives?
- Where do different learner groups experience barriers?

Such inquiry legitimises qualitative evidence, learner voice and practitioner judgement as essential components of understanding, rather than treating them as inferior to numerical indicators. On the basis of richer diagnostic understanding, organisations can then move into practice development and innovation, designing targeted interventions such as modified induction models, flexible assessment approaches, or enhanced pastoral support. At this stage, the taxonomy supports proportionality: claims are framed in terms of plausibility and feasibility rather than effectiveness, and learning is documented systematically. Evaluation becomes appropriate only once practices have stabilised and underlying theories of change are explicit. Mixed-method approaches allow organisations to test whether particular interventions are associated with improved engagement or progression for specific learner groups, while remaining alert to unintended consequences. Quantitative analysis retains an important role, but is interpreted alongside qualitative insight rather than treated as a stand-alone verdict.



APPLYING THE THREE-AXIS MODEL

The three-axis model offers valuable diagnostic perspective in this domain. Retention initiatives often show high levels of codification, through attendance policies, risk dashboards and reporting processes, while remaining relatively weak on rigour or transferability. The model helps leaders to identify whether codified practices are actually grounded in robust understanding, or whether they risk entrenching ineffective assumptions. Conversely, rich local inquiry may generate deep understanding of learner experience without being sufficiently codified to influence organisational practice. In such cases, leadership attention should focus on capturing, structuring and sharing learning rather than demanding further analysis. Retention and progression also raise ethical questions that cannot be resolved through data alone. Decisions about course withdrawal, assessment flexibility and continuation involve professional judgement, fairness and care.

The taxonomy explicitly recognises phronesis – practical wisdom – as an essential dimension of evidence-informed practice (Flyvbjerg, 2001). By situating retention work within a broader professional knowledge framework, the taxonomy guards against reductive interpretations of data and supports more ethically grounded decision-making. The final stage involves codifying learning from retention inquiry into guidance, standards and professional learning. This may include induction principles, learner support frameworks or progression heuristics that inform practice while remaining open to contextual adaptation. Professional bodies and networks play a crucial role in aggregating such learning, supporting transferability without imposing uniform solutions.

This second worked example demonstrates that the taxonomy is not limited to pedagogical or curriculum-focused improvement. It provides equally powerful support in domains where risk, accountability and learner vulnerability intersect, enabling organisations to balance rigour, care and professional agency in pursuit of sustained improvement.

SECTION 2: THEORETICAL FOUNDATIONS



A TAXONOMY OF RESEARCH-RELATED PRACTICE

This section proposes a continuum of research-related practice, from tacit reflection through to formal academic research and evidence synthesis. The intention is not to position these activities as a hierarchy of status, nor to imply that academic research is the only legitimate form of knowledge production. Instead, the taxonomy clarifies that different forms of research-related practice serve different purposes, require different levels of rigour, and produce different kinds of outputs. In FE and skills, these practices frequently overlap.

The same improvement project may contain reflective practice, structured inquiry, curriculum development, innovation, evaluation and dissemination. The value of the taxonomy lies in making these distinctions visible, enabling organisations to commission proportionately, develop capability intentionally, and create clearer pathways through which local learning becomes organisational and sector knowledge. The taxonomy is therefore best understood as a system of connected professional knowledge activities. It reflects the reality that FE and skills improvement is rarely linear. It is iterative, responsive and relational, shaped by learner diversity, local labour market needs, organisational capacity, and shifting policy expectations.

There are a number of underlying concepts that also merit note:



Professional knowledge and tacit expertise

Professional practice in FE and skills is inherently contextual, relational and adaptive. Much of what practitioners do well cannot easily be reduced to checklists or manuals. Polanyi's insight that professionals 'know more than they can tell' remains highly relevant (Polanyi, 1966). Eraut's work on professional learning demonstrates that expertise develops primarily through experience, informal learning and situational judgement rather than through formal instruction alone (Eraut, 1994; 2004). This challenges deficit narratives that frame FE practice as needing to import expertise from elsewhere, and instead reframes the task as one of surfacing, validating and extending existing knowledge.



Professionalism and democratic inquiry

Sachs (2016) distinguishes between managerial professionalism, driven by compliance and external accountability, and democratic professionalism, grounded in agency, ethical responsibility and collective improvement. Practitioner inquiry sits at the heart of democratic professionalism. In high-stakes accountability environments, inquiry enables practitioners to reclaim ownership of educational purpose, pedagogical judgement and quality (Ball, 2015). It allows values, not just metrics, to inform improvement decisions.



Systems learning

From a systems perspective, sustainable improvement requires more than scaling interventions. It depends on feedback loops, shared meaning and infrastructure for collective learning (Senge, 2006). Research operates as distributed intelligence, enabling local insight to become system knowledge when shared and codified through networks, intermediaries and professional bodies (Greany and Higham, 2018).

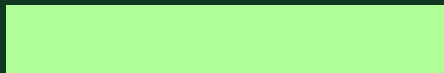


SECTION 3:

A FIVE-TIER TAXONOMY

OF RESEARCH-RELATED

PRACTICE



For leadership audiences, the taxonomy is presented here in a simplified five-tier form. This model clarifies the progression from tacit professional learning to codified, transferable knowledge. For readers who want additional precision without the full methodological detail, Appendix A can also be used as a reference list of Levels 1–14, with each level providing a clear label for common forms of FE and skills inquiry and evaluation.

The five tiers represent increasing levels of rigour, transferability and codification, while recognising that movement between them is not strictly linear.

Tier 1:

Professional sense-making

This tier encompasses reflective practice, professional dialogue and communities of practice. Learning is primarily tacit and localised. Practitioners interpret experience, adapt practice and share insights informally. While essential to professional growth, knowledge at this stage remains fragile unless captured or tested.

Tier 2:

Structured inquiry and evidence-informed improvement

At this tier, practice questions become structured. Professional inquiry, evidence-informed decision-making and quality improvement activity begin to generate explicit local knowledge. Evidence is gathered systematically, though transferability remains limited. This tier is critical for moving beyond anecdote.

Tier 3:

Innovation and development

This tier focuses on the design and refinement of new approaches. Practice development, curriculum innovation and experimentation occur here. Feasibility and early outcomes are tested. The emphasis is on creation and refinement rather than proof of impact.

Tier 4:: Testing, evaluation and validation

Here, interventions are tested across contexts. Pilots, trials and evaluation establish effectiveness, implementation conditions and mechanisms of change. Transferability becomes a central concern. Evidence at this stage supports defensible organisational and system decisions.

Tier 5:: Knowledge production and mobilisation

The final tier encompasses practitioner research, action research, applied research, academic research and evidence synthesis. Knowledge becomes codified into frameworks, standards, toolkits and sector guidance. At this stage, research functions as system infrastructure rather than isolated activity.

The five-tier model provides a strategic overview of how research-related practice develops in FE and skills. The detailed taxonomy in Appendix A elaborates the specific forms of activity within each tier, enabling organisations to design proportionate research and improvement strategies aligned to their capacity and ambition.



CROSS-CUTTING CAPABILITIES

In practice, many of the most useful FE and skills studies are mixed-method by design, combining quantitative pattern recognition with qualitative explanation. Quantitative analysis is best understood as a methodological capability rather than a discrete category of practice. Across the taxonomy, quantitative methods support reflection, monitoring, evaluation and research, ranging from descriptive statistics to inferential analysis and modelling (Cohen et al., 2018). Similarly, review spans multiple levels, from local performance review to rapid evidence reviews for innovation, through to systematic review and meta-analysis capable of informing sector-wide guidance (Petticrew and Roberts, 2006). Understanding these as cross-cutting capabilities avoids false binaries between quantitative and qualitative or between review and research and instead foregrounds fitness for purpose.

Equally, action research and quality improvement are often closely related in practice, but they are not identical. Quality improvement is primarily outcomes-driven and operational, whereas action research is explicitly knowledge-generating: it seeks to produce learning that is sufficiently explicit, evidenced and shareable to inform wider practice. In FE and skills, the most effective improvement cultures are those where quality improvement is underpinned by inquiry, and where action research provides a disciplined structure for testing assumptions rather than simply responding to performance signals.

This distinction is important because evaluation is often treated as a compliance mechanism rather than a learning mechanism. Patton's utilisation-focused evaluation framework reinforces that the value of evaluation lies not only in methodological design but in the extent to which findings are actually used to inform decisions, build capability and strengthen improvement practice (Patton, 2008). In FE and skills, where reform cycles are frequent and implementation conditions vary significantly, evaluation must therefore be designed explicitly for learning and mobilisation, not simply for accountability reporting.

EXAMPLE 3

Ofsted identifies professional learning and development as a priority theme, reflecting an increasing expectation that workforce capability is a central driver of quality and learner outcomes (Ofsted, 2026). This theme maps directly onto the five-tier taxonomy, demonstrating how professional learning activity can evolve from informal reflection into codified, transferable system knowledge.

At Tier 1 (**professional sense-making**), professional learning begins through reflective practice and professional dialogue. Teachers notice patterns in learner engagement, assessment outcomes, and behaviour, and adjust practice in real time. This learning is highly contextual and often tacit, but it represents the foundational layer of professional knowledge creation.

At Tier 2 (**structured inquiry and evidence-informed improvement**), organisations begin to test CPD approaches systematically. For example, a provider may investigate whether instructional coaching improves learner progress in English and maths, or whether subject mentoring strengthens retention among early-career teachers. Evidence is gathered through learner outcome analysis, structured observation, staff feedback and learner voice.

At Tier 3 (**innovation and development**), providers design and refine new workforce development models. This may include digital CPD platforms, peer-led scholarship programmes, micro-credential approaches, or cross-provider mentoring networks. Early experimentation focuses on feasibility, participation and implementation quality.

At Tier 4 (**testing, evaluation and validation**), approaches are evaluated more rigorously across departments, sites or cohorts. At this stage, the key question shifts from ‘is this promising?’ to ‘does it work reliably, for whom, and under what conditions?’. Mixed-method evaluation supports both impact assessment and implementation learning.

At Tier 5 (**knowledge production and mobilisation**), professional learning becomes codified and transferable. Practitioner research and action research generate publishable case studies, tested intervention models, and practical guidance. Applied research can consolidate findings across providers, and evidence synthesis can translate emerging practice into sector-wide CPD frameworks, professional standards and improvement toolkits.

This example illustrates how professional learning and development can travel through the taxonomy: beginning with reflective practice, strengthening through structured inquiry and evaluation, and ultimately becoming codified knowledge that supports sustained professionalism and system improvement. (See Appendix A for the detailed level mapping and expanded worked Example A1.)



SUMMARY: HOW THE TAXONOMY SHOULD BE USED

The summary below consolidates the tier model into a commissioning lens, showing how activity should mature across time rather than being treated as a single event. The taxonomy should be used as a shared organisational language for describing activity and designing proportionate research and improvement strategies. It enables leaders to identify where current activity sits, what kind of evidence is required, and what kind of infrastructure is needed to move learning forward. Crucially, the taxonomy also makes visible that improvement does not end with evidence generation. The ultimate goal is codification and mobilisation: converting learning into professional tools, frameworks and practices that can be sustained, adapted and scaled across FE and skills.

The taxonomy positions research-related activity along a continuum from reflective practice to evidence synthesis, without implying a hierarchy of value. Each form of activity serves different purposes, audiences and standards of evidence, and each is necessary to a functioning professional system.

1

At the foundational level, reflective practice and professional dialogue support immediate sense-making and adaptation (Schön, 1983; Wenger, 1998).

2

Structured reflection enables practitioners to interrogate assumptions and respond intelligently to complexity.

3

Professional inquiry and quality improvement introduce intentionality and structure. They generate local evidence to inform decisions, test approaches and develop early models of practice (Stenhouse, 1975; Argyris and Schön, 1978).

4

Practice development and innovation involve deliberate design and experimentation. At this stage, ideas are promising but provisional. Evaluation and piloting add rigour, addressing questions of effectiveness, feasibility and transferability through theory-informed and mixed-method approaches (Pawson and Tilley, 1997).

5

At the upper end of the continuum, practitioner research, applied research and academic research generate publishable and generalisable knowledge, while evidence synthesis consolidates findings into forms capable of shaping standards, frameworks and professional learning (Flyvbjerg, 2006; Petticrew and Roberts, 2006).

SECTION 4:

THE THREE-AXIS MODEL OF RIGOUR, TRANSFERABILITY AND CODIFICATION

A three-axis model of rigour, transferability and codification is the conceptual core of the framework. It is designed to move organisations away from linear or hierarchical conceptions of research quality and towards a more accurate representation of how professional knowledge is generated, validated and mobilised within complex systems such as FE and skills. This supports the signals from Ofsted (2026) that sector-wide evidence needs to extend beyond isolated studies, underscoring the importance of building knowledge that is transferable across contexts and usable by providers, regulators and policymakers alike.

The three-axis model proposed here positions research-related activity according to rigour (methodological strength), transferability (applicability across contexts), and codification (explicitness and teachability). The taxonomy describes the range of activity; the three-axis model provides a way of diagnosing the strength, maturity and system value of that activity.



Figure 1: The three-axis model of rigour, transferability and codification





HOW TO INTERPRET FIGURE 1

The triangle should be read diagnostically rather than judgementally. Activity located towards the lower centre is often highly contextually relevant but weakly transferable and insufficiently codified. Activity towards the left is increasingly rigorous and defensible. Activity towards the right is increasingly embedded as professional infrastructure. Activity towards the top is increasingly system-relevant and capable of informing improvement beyond the originating organisation. The value of the model lies in revealing imbalance. The strategic question is not whether this activity is high quality, but where is this activity strong, where is it weak, and what kind of investment is needed to strengthen it?

Traditional hierarchies of research quality, often implicitly modelled on experimental or academic paradigms, assume that knowledge progresses in a linear fashion from low-value local insight to high-value generalisable evidence. In professional systems, this assumption is deeply misleading. It obscures the fact that highly rigorous research may have little practical influence, while locally powerful practices may be widely adopted despite weak evidential foundations. Within FE and skills, linear hierarchies also risk delegitimising practitioner knowledge, reinforcing perceptions that only externally produced research counts. The three-axis model responds to this problem by recognising that different qualities of knowledge matter simultaneously, and that imbalance across these dimensions is often the root cause of ineffective improvement.

The model is therefore not a ladder of status. It is a diagnostic tool for strengthening organisational learning and increasing the likelihood that professional knowledge becomes sustainable, transferable and usable at scale.



THE THREE DIMENSIONS EXPLAINED

Rigour

refers to the methodological strength, validity and transparency of inquiry. It concerns the appropriateness of design, the coherence between questions and methods, and the trustworthiness of conclusions (Lincoln and Guba, 1985; Cohen et al., 2018). Rigour is not synonymous with complexity. In early-stage inquiry, rigour may be demonstrated through careful documentation, reflexivity and ethical practice rather than formal research design.

Transferability

relates to the extent to which learning can inform practice beyond the immediate context in which it was generated. In professional systems, transferability depends less on statistical generalisation and more on thick description: clear articulation of context, assumptions and mechanisms that allow others to judge relevance (Lincoln and Guba, 1985). It is important to distinguish transferability from statistical generalisability. In FE and skills, where learner cohorts, labour markets and institutional conditions vary significantly, the most useful evidence is often evidence that makes its assumptions and mechanisms visible, enabling other organisations to judge relevance intelligently rather than treating findings as universally applicable.

Codification

concerns the degree to which knowledge is rendered explicit, documented and teachable. Codified knowledge can be embedded in curricula, standards, CPD frameworks, professional guidance and organisational routines. Without codification, learning remains fragile and person-dependent; with codification alone, it risks becoming ritualised and detached from underlying purpose (Nonaka and Takeuchi, 1995). In FE and skills, codification often takes the form of curriculum sequencing principles, teaching and learning playbooks, assessment and feedback models, induction frameworks, learner support protocols, mentoring structures, and coaching resources. These artefacts represent the point at which inquiry becomes organisational memory: practice knowledge that can be taught, replicated, adapted and sustained beyond individual staff members.



TYPICAL IMBALANCE PATTERNS IN FE AND SKILLS

Mapping activity across these three dimensions reveals common imbalance patterns.

First

high codification / low rigour is frequently observed in compliance-driven environments.

Policies, frameworks and tools proliferate but are weakly grounded in systematic inquiry. Such practices are stable but not necessarily effective and can entrench poor assumptions.

Second

high rigour / low transferability characterises much externally commissioned research.

Findings may be methodologically strong, but are insufficiently contextualised or translated into usable forms, limiting their impact on practice.

Third

high insight / low codification is typical of practitioner-led inquiry.

Rich understanding is generated but remains local and vulnerable to staff turnover or shifting priorities. In these cases, the problem is not lack of rigour, but lack of deliberate knowledge mobilisation.

The diagnostic power of the model lies in making these patterns visible and actionable.

For leaders and boards, the three-axis model provides a practical tool for:

- ☒ **commissioning inquiry proportionately to purpose**
- ☒ **identifying where investment should focus (methods, dissemination or infrastructure)**
- ☒ **avoiding unrealistic expectations of early-stage innovation**
- ☒ **assuring that codified practice is grounded in credible learning.**

Crucially, the model does not prescribe an ideal 'high-high-high' endpoint for all activity. Instead, it supports intentional movement across dimensions over time, aligned to organisational maturity and purpose.



IMPLICATIONS FOR COMMISSIONING AND EVALUATION

When commissioning research or evaluation, the model supports clearer articulation of what is required. Early-stage initiatives may prioritise transferability and codification over rigour, while high-stakes policy pilots may require rigorous design and explicit mobilisation strategies. Making these choices explicit increases confidence, credibility and learning value.



KNOWLEDGE CREATION AND PROFESSIONAL LEARNING

The three-axis model aligns with theories of organisational knowledge creation, where learning moves through cycles of socialisation, externalisation, combination and internalisation (Nonaka and Takeuchi, 1995). Codification is not the end point of learning, but a transitional phase that supports collective sense-making and further adaptation. In FE and skills, professional learning depends on maintaining movement between tacit and explicit knowledge, between local insight and shared standards. The model provides a conceptual scaffold for supporting that movement. In doing so, it also makes visible a critical distinction in research value: between knowledge that remains abstracted from practice, and knowledge that is demonstrably capable of informing professional judgement, improving provision and shaping learner outcomes.



ETHICAL AND PROFESSIONAL CONSIDERATIONS

Finally, the model reinforces the ethical dimension of evidence-informed practice and explicitly surfaces the relative value of research that is grounded in practice, application and impact over research that is largely esoteric or detached from use. The framework does not dismiss theoretical or discipline-building work; rather, it challenges the sector to be honest about what kinds of knowledge are needed for improvement, by whom, and for what purpose. This matters because evidence is not neutral in high-stakes systems. In accountability-driven environments, data and research findings can be used to legitimise predetermined decisions, reinforce deficit narratives about learners, or narrow professional judgement into compliance routines. A research-engaged sector therefore requires not only methodological capability, but ethical literacy: the capacity to interpret evidence responsibly, to recognise what is missing from datasets, and to protect the centrality of learner experience, inclusion and professional care.

In FE and skills, value lies not simply in methodological sophistication, but in the productive relationship between inquiry and use. Research that is rigorous yet disconnected from practice may advance academic debate, but it contributes little to sector learning unless it is rendered transferable and codified into usable forms. Conversely, practice-anchored inquiry that generates credible local learning and demonstrable improvement outcomes should be recognised as legitimate and valuable knowledge production, even where it does not conform to conventional academic norms. By making rigour, transferability and codification visible as distinct dimensions, the model legitimises impact-oriented, practice-engaged research and exposes the limitations of work that is high in abstraction but low in application. This supports more responsible use of evidence in contexts where decisions have profound consequences for learners, avoiding the marginalisation of professional judgement and learner voice in favour of technically elegant but context-poor evidence.

This framework is not presented as a definitive classification of all research practice, nor does it assume a single pathway of progression. Its purpose is heuristic: to support clearer commissioning, stronger professional ownership, and more deliberate movement from local insight to codified system learning.

SECTION 5: ORGANISATIONAL IMPLICATIONS: LEADERSHIP, GOVERNANCE AND SYSTEM STEWARDSHIP



KNOWLEDGE BUILDING: FROM DATA TO WISDOM

The framework has significant implications for leadership practice, governance arrangements and the role of professional bodies within FE and skills. Aligning research practice with external statements of evidence need, such as Ofsted's Areas of Research Interest (2026), can support better system learning, reduce duplication, and ensure that research investment is relevant to inspection, regulatory and policy priorities. Crucially, this reframes research not as a marginal or specialist function but as a core component of organisational and system capacity. In understanding how research and inquiry operate as system infrastructure, it is helpful to distinguish between different forms of knowing.

A widely used model within knowledge management is the DIKW hierarchy, which describes how understanding develops through four linked stages: data, information, knowledge and wisdom (Rowley, 2007). The value of this model is that it clarifies why FE and skills can be rich in performance metrics while still struggling to generate sustained improvement.

- **Data** refers to raw signals and measures: attendance rates, retention figures, achievement outcomes, learner survey responses, employer feedback and behavioural logs. In isolation, these data points do not explain anything. They are simply indicators that something is happening.
- **Information** is created when data is organised and interpreted into patterns, summaries and comparisons. Dashboards, cohort analyses, trend reports and inspection self-assessment narratives all represent information. Information can describe what is changing and where, but it does not necessarily reveal why.

- **Knowledge** emerges when information is combined with professional interpretation, inquiry and evidence, allowing organisations to develop explanatory understanding. It strengthens as insight becomes explicit, tested and shared. Wisdom emerges when evidence is integrated with ethical and professional judgement, particularly where improvement intersects with inclusion and wellbeing (Flyvbjerg, 2001). In FE and skills this might mean recognising that declining attendance is not merely a behavioural issue, but is linked to timetabling, transport barriers, mental health, confidence, curriculum relevance, or learner identity. Knowledge is therefore not simply the accumulation of more data, but the development of insight that supports action.
- **Wisdom** represents a further shift. It is the capacity to apply knowledge with judgement, ethics and contextual sensitivity. In professional systems such as FE and skills, wisdom is the point at which leaders and practitioners ask not only what works, but what should be done, for whom, and at what cost or risk.

This is particularly important in a sector where many learners have experienced prior disadvantage and where improvement strategies can have unintended consequences, including exclusionary practices or reduced trust. Further, inspection frameworks increasingly expect providers to demonstrate evidence of improvement that goes beyond surface metrics, aligning with the view that data is a starting point for inquiry rather than an end in itself.

The DIKW hierarchy, therefore, reinforces a central argument of this report: research capacity is not primarily about generating more data, but about creating the infrastructure through which data becomes usable knowledge and, ultimately, professional wisdom capable of sustaining improvement over time.



LEADERSHIP AND EXECUTIVE DECISION-MAKING

For executive leaders, the taxonomy and three-axis model provide a disciplined way to think about what kind of learning is required, for what purpose, and with what expectations. Rather than defaulting to generic calls for 'evidence-based practice', leaders can make more nuanced decisions about when professional inquiry, innovation, evaluation or research is appropriate.

This enables leaders to:

- commission inquiry proportionately and credibly
- protect early-stage innovation from premature judgement
- articulate realistic expectations of impact and timescale
- balance improvement ambition with professional confidence and morale.

Importantly, the framework supports leaders to resist the false binary between action and analysis. Inquiry becomes an integral part of improvement, not a precursor or an optional add-on.



GOVERNANCE, ASSURANCE AND BOARD OVERSIGHT

At board level, the framework offers a structured basis for assurance. Boards are frequently required to make judgements about organisational effectiveness, learner outcomes and responsiveness to external challenge, often on the basis of data and risk indicators alone. The taxonomy provides additional lenses through which boards can test the robustness of improvement narratives. Boards can use the framework to ask more intelligent questions, such as the following questions.

- What forms of inquiry underpin this strategy or intervention?
- How is learning from practice being evaluated and shared?
- Where does this activity sit in terms of rigour, transferability and codification?
- What risks arise from over- or under-reliance on particular forms of evidence?

In this sense, research capacity becomes part of corporate assurance, strengthening organisational self-knowledge and reducing over-dependence on external judgement.



BUILDING ORGANISATIONAL CAPABILITY

Embedding the framework requires deliberate investment in capability. This does not imply that all staff become researchers, but that organisations cultivate:

- research literacy among leaders and middle managers
- confidence in using mixed forms of evidence
- skills in documentation, reflection and dissemination
- ethical awareness in evidence-informed decision-making.

Such capability development strengthens professional capital across the organisation and supports more resilient improvement cultures over time (Hargreaves and Fullan, 2012).



THE ROLE OF ETF

As the professional body for the FE and skills sector, ETF occupies a critical position within the knowledge ecosystem of FE and skills. ETF is the intermediary between local practice and system learning, translating distributed inquiry into shared standards, guidance and identity.

Within this framework, ETF is uniquely placed to:

- support common language and expectations about research and inquiry
- curate and synthesise practitioner-generated evidence
- embed learning within professional standards, CPD and career pathways
- legitimise diverse forms of professional knowledge alongside academic research.

This intermediary role is especially important in a fragmented system where no single institution holds responsibility for collective learning. The framework, therefore, strengthens the case for ETF as a steward of knowledge, not merely a provider of training or accreditation.



IMPLICATIONS FOR KEY ACTORS ACROSS FE AND SKILLS

A central strength of the taxonomy is that it clarifies roles across the system. Research capacity cannot be built by providers alone, nor can it be outsourced entirely to universities or external agencies. Sustainable system learning requires shared responsibility.

- 1** For FE and skills providers, the framework supports clearer commissioning and more proportionate evaluation design. It enables organisations to invest in professional inquiry as part of routine improvement and to strengthen codification so that learning is retained and shared.
- 2** For ETF, the framework identifies a key role as a steward of knowledge, where the contribution lies in curating evidence, supporting professional standards, convening communities of practice, and translating practitioner inquiry into shared tools and CPD.

- 3 For policymakers and funders, the framework provides a basis for more realistic expectations of evidence. It reinforces the importance of investing in research infrastructure and dissemination rather than commissioning episodic studies that have limited sustained impact.
- 4 For universities and research partners, the framework strengthens the case for partnership models grounded in co-production. It positions practitioners not as objects of study, but as knowledge producers, and encourages methodological approaches that respect context while building transferability. In practice, usable evidence looks different for different audiences: for practitioners it may be guidance, exemplars and tools that can be applied in classrooms; for middle leaders it may be tested models and implementation lessons; for senior leaders and boards it may be evaluation findings, risk signals and decision-ready insights; and for policymakers it may be synthesis that clarifies what can transfer across contexts.



SYSTEM LEARNING AND LONG-TERM IMPROVEMENT

Finally, the framework contributes to a vision of FE and skills as a learning system rather than a delivery mechanism. By valuing the full continuum of inquiry and strengthening pathways between local insight and shared knowledge, the sector can move beyond episodic reform towards cumulative improvement. This requires patience, trust and sustained leadership. However, the alternative of a short-cycle reform driven by metrics without learning, risks perpetuating instability and professional disengagement. The taxonomy and three-axis model offer practical tools for supporting a more mature, learning-led system.



CONCLUSION

This report reframes FE and skills as a knowledge-producing professional system, positioning research not as an external or discretionary activity but as core infrastructure for improvement. By valuing the full continuum of research-related practice and strengthening the pathways between reflective practice, inquiry, innovation, evaluation and evidence synthesis, the sector can move beyond episodic reform towards sustained, ethical and evidence-informed professionalism. The three-axis model of rigour, transferability and codification provides leaders, professional bodies and policymakers with a practical diagnostic tool, making visible where improvement activity is strong or fragile and where investment is required in methodological capability, dissemination and codification. The report's central argument is that FE and skills will not become a self-improving system through compliance alone, but through deliberate inquiry cultures and the infrastructures needed to mobilise learning at scale.

APPENDIX A

THE DETAILED

TAXONOMY



This appendix sets out the full taxonomy underpinning the five-tier model presented in Section 4. This detailed taxonomy is provided to support practitioners, evaluators and research partners who require greater specificity about methods, outputs and expectations. It also enables providers to map existing activity more precisely and to identify what infrastructure is needed to move work from local insight to codified system knowledge.

LEVEL 1: REFLECTIVE PRACTICE (INDIVIDUAL SENSE-MAKING)

Reflective practice refers to the continuous professional process of learning from experience. Schön (1983) distinguishes reflection-in-action (adaptation during practice) and reflection-on-action (post-event analysis). Reflective practice is central to teaching and leadership effectiveness, particularly in FE and skills contexts where practitioners must respond to diverse learner needs, complex barriers to participation, and rapidly changing curriculum and assessment requirements. Reflective practice is typically informal and unstructured.

It may involve lesson review, reflective journaling, professional supervision, or simply structured thinking after a teaching or leadership event. Its value lies in the way it develops professional judgement and adaptive expertise over time. However, its limitations are also clear: reflective practice is often tacit, rarely recorded, and therefore difficult to transfer beyond the individual practitioner. Typical outputs include personal insight, small adjustments to teaching practice, improved learner interactions, and strengthened professional confidence. In knowledge terms, reflective practice generates early hypotheses and emerging explanations, but these remain fragile unless shared, tested or codified.

LEVEL 2: PROFESSIONAL DIALOGUE AND COMMUNITIES OF PRACTICE (COLLECTIVE REFLECTION)

At this stage, reflective learning becomes social. Professional dialogue allows practitioners to test their interpretations, share emerging insights and build shared language about what works, for whom, and under what conditions. Wenger (1998) argues that communities of practice are central to professional learning, enabling groups to develop shared repertoires of practice and shared professional identity. In FE and skills, communities of practice often emerge through curriculum teams, coaching networks, peer observation processes, subject specialist groups, and professional development communities.

These spaces are critical because they begin to convert individual tacit knowledge into collective professional understanding. This also aligns with Freire's emphasis on the relationship between reflection and action as a mechanism of liberation and learning. Through collective dialogue and shared inquiry, practitioners do not merely exchange practice tips; they generate new professional knowledge through a process of reflection, action and re-reflection. In this sense, both teachers and learners can be understood as co-creators of knowledge, continuously reshaping practice and reality through shared meaning-making (Freire, 1970). This is particularly resonant in FE and skills, where learner experience, identity and agency are often central to understanding what works. Typical outputs include shared pedagogical approaches, informal practice norms, peer learning, and early articulation of improvement priorities. However, this stage remains vulnerable to becoming anecdotal unless supported by evidence gathering or structured inquiry.

LEVEL 3: EVIDENCE-INFORMED PRACTICE (USING EXTERNAL EVIDENCE TO SHAPE LOCAL DECISION-MAKING)

Evidence-informed practice refers to the use of existing research evidence, evaluation findings, or sector guidance to shape decisions about curriculum, pedagogy, learner support or organisational improvement. In this context, practitioners and leaders are not necessarily producing new evidence, but they are acting as informed consumers of evidence. This stage is increasingly important as the FE and skills sector seeks to develop a more robust relationship with the wider evidence base. However, Biesta (2010) cautions against simplistic what works approaches that treat evidence as universally transferable.

Evidence-informed practice is strongest when it is rooted in professional judgement, contextual understanding and ethical reasoning, rather than being treated as a technical exercise of implementation. Typical outputs include improved intervention selection, stronger confidence in curriculum and pedagogical decisions, and enhanced organisational capacity to interpret research claims critically. Evidence-informed practice can also act as a bridge into professional inquiry and as a knowledge production mechanism, where external evidence is tested and adapted within local settings.

LEVEL 4: PROFESSIONAL INQUIRY (STRUCTURED PRACTICE QUESTIONING):

Professional inquiry refers to structured investigation of practice-based questions, typically oriented towards improvement in a local setting. It may be conducted at individual, departmental or organisational level and often emerges from reflective practice and professional dialogue. Professional inquiry is particularly significant in FE and skills because it supports practitioners to move beyond anecdotal improvement narratives and towards disciplined evidence-building. Stenhouse (1975) is foundational here, positioning teachers as curriculum researchers who investigate their own practice systematically.

In FE and skills, professional inquiry frequently focuses on questions of engagement, attendance, retention, learner motivation, inclusion, progression and employability outcomes. Professional inquiry typically involves manageable forms of evidence gathering, such as structured learner voice, observation, simple cohort comparisons, analysis of internal performance data, and small-scale qualitative analysis. Typical outputs include locally valid evidence, strengthened improvement planning, and early models of what appears to work in specific contexts. Professional inquiry often produces findings that are credible enough to shape organisational decisions, but not yet robust enough to claim wider transferability.

LEVEL 5: QUALITY IMPROVEMENT (SOLUTIONS-DRIVEN ITERATIVE IMPROVEMENT):

Quality improvement differs from research because it is explicitly designed to improve outcomes rapidly rather than to generate generalisable knowledge. It is often cyclical and iterative, using approaches such as Plan–Do–Study–Act cycles, process mapping, root cause analysis and performance dashboards. In FE and skills, quality improvement is a core operational necessity. Organisations must continuously respond to participation patterns, achievement gaps, learner support needs, curriculum compliance and inspection pressures.

Quality improvement provides the practical tools through which organisations respond to these demands. However, the limitations of quality improvement are also important. Argyris and Schön (1978) distinguish between single-loop learning (adjusting actions to meet targets) and double-loop learning (challenging the underlying assumptions and values shaping decisions). Quality improvement can become compliance-driven if it focuses only on performance metrics without deeper inquiry into causality, learner experience, or structural barriers. Typical outputs include measurable improvements in performance indicators, refined organisational processes, and operational learning. The challenge is that this learning is often not codified or disseminated beyond the immediate team, meaning improvement remains localised and fragile.

LEVEL 6: PRACTICE DEVELOPMENT AND CURRICULUM DESIGN (STRUCTURED CREATION)

Practice development refers to the intentional development of improved teaching models, curriculum structures, learner support processes, or organisational approaches. It is typically driven by professional insight and informed by evidence but is focused on design and creation rather than testing. In FE and skills, practice development is a key expression of sector expertise. It is where organisations design provision that reflects learner needs, employer demand, place-based economic priorities, and the distinctive pedagogical requirements of vocational and technical education. Practice development often involves co-design with employers, curriculum sequencing, development of assessment strategies, design of blended delivery models, and creation of learner support frameworks. It may also include professional standards work and staff development programme design. Typical outputs include schemes of work, curriculum frameworks, teaching resources, learner support interventions and staff development materials. These outputs are often highly codified, but their effectiveness may not yet be evidenced unless linked to evaluation.

LEVEL 7: INNOVATION (NEW APPROACHES, NEW TOOLS, NEW MODELS)

Innovation refers to the deliberate development of novel solutions, often in response to new policy contexts, new technologies, changing learner needs, or emerging labour market demands. Innovation may occur in pedagogy, curriculum design, digital learning approaches, learner support, employer engagement, or organisational operating models. Innovation is essential in FE and skills because the sector operates in a high-change environment. However, innovation is frequently constrained by limited resource, high accountability and low organisational slack. As a result, innovation can become fragmented, short-term, and insufficiently evaluated. Typical outputs include prototypes, new delivery models, new pedagogical approaches, digital tools, and emerging practice frameworks. At this stage, evidence is often based on early indicators and practitioner judgement rather than validated outcomes.



LEVEL 8: EXPERIMENTATION AND FEASIBILITY TESTING (PROOF OF CONCEPT)

Experimentation is the early testing stage where innovations are assessed for feasibility, acceptability and early outcome signals. It asks whether an innovation appears to work sufficiently well to justify further investment, scaling or evaluation. In FE and skills, experimentation is often pragmatic. It may involve small pilots, rapid feedback loops, early learner voice, short pre/post comparisons, or informal comparisons across cohorts. The focus is often on identifying what is workable in real delivery contexts, rather than proving impact definitively. Typical outputs include refined intervention models, clearer delivery requirements, early evidence of promise, and go/no-go decisions. This stage is critical because it prevents premature scaling of interventions that are not feasible or not aligned with learner needs.

LEVEL 9: PILOTS AND TRIALS (TESTING EFFECTIVENESS ACROSS SETTINGS)

Experimentation is the early testing stage where innovations are assessed for feasibility, acceptability and early outcome signals. It asks whether an innovation appears to work sufficiently well to justify further investment, scaling or evaluation. In FE and skills, experimentation is often pragmatic. It may involve small pilots, rapid feedback loops, early learner voice, short pre/post comparisons, or informal comparisons across cohorts. The focus is often on identifying what is workable in real delivery contexts, rather than proving impact definitively. Typical outputs include refined intervention models, clearer delivery requirements, early evidence of promise, and go/no-go decisions. This stage is critical because it prevents premature scaling of interventions that are not feasible or not aligned with learner needs.

LEVEL 10: MONITORING, EVALUATION AND IMPACT ASSESSMENT

Evaluation seeks to understand not only whether outcomes changed, but why they changed, for whom, and under what conditions. Pawson and Tilley's (1997) realist evaluation tradition is highly relevant here, particularly in complex systems such as FE and skills where causality is rarely linear. Evaluation may include theory of change approaches, contribution analysis, mixed-method impact assessment and cost-effectiveness analysis. It may examine learner outcomes, organisational outcomes, workforce outcomes, and system outcomes.

In FE and skills, evaluation is particularly important because interventions may appear successful in performance terms while producing unintended consequences, such as exclusionary practice, narrowing of curriculum, or hidden disengagement. Evaluation therefore supports ethical accountability as well as performance improvement.

Typical outputs include credible impact claims, refined intervention design, learning about implementation, and evidence to support scaling or discontinuation.

LEVEL 11: PRACTITIONER RESEARCH AND ACTION RESEARCH

Practitioner research refers to research undertaken by practitioners within professional contexts, often with explicit methodologies, ethical protocols, and dissemination beyond the organisation. Practitioner research is not restricted to one methodological tradition. It may include action research, case study research, ethnography, practitioner-led evaluation, participatory inquiry, or mixed-method inquiry approaches. Action research has particular significance in FE and skills. Originating in Lewin's (1946) cyclical model of planning, acting, observing and reflecting, action research has developed into a participatory tradition in which inquiry is conducted with practitioners and communities as co-producers of knowledge (Reason and Bradbury, 2008). Kemmis, McTaggart and Nixon (2014) further emphasise action research as a critical participatory practice oriented towards transformation and professional agency. Action research is especially relevant because it functions as a bridging methodology across the continuum. It connects reflective practice, professional inquiry, innovation, implementation and evaluation into an iterative cycle.

In this sense, action research should not be treated as a discrete tier of activity, but as a methodological engine through which practice becomes professional knowledge. Typical outputs include publishable case studies, validated improvement learning, codified teaching resources, implementation guidance, and strengthened professional identity. Practitioner research strengthens not only organisational improvement but also sector legitimacy, making visible the distinctive pedagogies and relational expertise of FE and skills practice. Its wider system value emerges when findings are codified into usable artefacts and shared through networks, professional standards and sector communities.

LEVEL 12: APPLIED RESEARCH (SYSTEM-ORIENTED, DECISION-RELEVANT EVIDENCE)

Applied research seeks to generate transferable evidence that can inform system decisions. It often sits between practitioner research and academic research, combining methodological rigour with practical relevance.

Applied research may be commissioned by sector bodies, funders or government, and is often intended to inform policy development, programme design or system reform. Applied research may involve large-scale surveys, longitudinal analysis, comparative studies across multiple providers, workforce modelling, labour market analysis, and evaluation of national programmes. Typical outputs include scalable evidence, policy-relevant recommendations, system-level insight and implementation learning. Applied research is particularly important in FE and skills because many improvement challenges are structural rather than purely pedagogical, including workforce supply, funding incentives, qualification design, and place-based labour market alignment.

LEVEL 13: ACADEMIC RESEARCH (THEORY-BUILDING AND DISCIPLINARY CONTRIBUTION)

Academic research contributes to conceptual understanding and theory-building. It may examine pedagogy, vocational identity, curriculum theory, professionalisation, governance, funding models, learner transitions and policy dynamics. Academic research is essential because it provides the intellectual infrastructure through which the sector can articulate its distinctiveness and challenge deficit narratives. Academic research often involves peer review, methodological scrutiny, and theoretical development. It may use quantitative, qualitative or mixed-method approaches and may extend beyond FE and skills into broader vocational education and training (VET/TVET) scholarship. Typical outputs include peer-reviewed journal articles, conceptual frameworks, doctoral research and disciplinary contribution. Academic research strengthens the long-term knowledge base of the sector and supports FE and skills to be understood as a legitimate field of scholarship rather than solely a delivery system.

LEVEL 14: EVIDENCE SYNTHESIS AND KNOWLEDGE MOBILISATION

Evidence synthesis consolidates findings across multiple studies, building stronger conclusions than any single project can provide. Petticrew and Roberts (2006) describe systematic review as a method for strengthening evidence claims through transparent synthesis processes. Meta-analysis can provide quantitative aggregation of effect, while narrative synthesis can integrate qualitative and mixed-method studies to develop richer explanatory understanding. Evidence synthesis and knowledge mobilisation are particularly significant because they are the stages at which research becomes sector infrastructure. Without synthesis and mobilisation, research remains fragmented and episodic. Without translation, evidence remains inaccessible to practitioners and leaders. Typical outputs include systematic reviews, evidence toolkits, professional guidance, research-informed CPD programmes, professional standards, and sector frameworks. These outputs represent the point at which knowledge becomes transferable and durable: not simply published, but usable.

EXAMPLE A1

Ofsted's Areas of Research Interest identifies professional learning and development as a priority theme, reflecting an increasing expectation that workforce capability is not only a delivery issue but a central driver of quality and learner outcomes (Ofsted, 2026). This theme maps directly onto the proposed taxonomy, demonstrating how everyday practice-based learning can be developed into codified, transferable sector knowledge.

AT LEVEL 1

(reflective practice), professional learning begins informally. Individual teachers reflect on lesson effectiveness, learner engagement, behavioural dynamics, or assessment practice. This reflection may be prompted by routine performance data, learner feedback, or mentoring conversations. The knowledge generated at this stage is largely tacit: it is highly valuable, but fragile and difficult to transfer.

AT LEVEL 2

(professional dialogue and communities of practice), these reflections become collective. Teaching teams share strategies, compare learner responses, and develop informal consensus about what appears to work. This is often the stage at which effective practice begins to emerge as shared language, though it may remain anecdotal unless supported by inquiry.

AT LEVEL 3

(evidence-informed practice), professional learning becomes more explicit and intentional. Teams draw on external evidence, such as research on coaching, cognitive science, vocational pedagogy, or inclusive teaching practice. This supports improved decision-making about CPD priorities and reduces reliance on fashionable or untested interventions.

AT LEVEL 4

(professional inquiry), the organisation begins to test CPD approaches systematically. For example, a provider might ask whether instructional coaching improves learner progress in English and maths, or whether subject-specialist mentoring increases retention among new staff. Evidence may include learner outcomes, staff confidence measures, structured observation, or learner voice.

AT LEVEL 5

(quality improvement), findings are translated into operational change. The provider may redesign CPD time, introduce structured coaching cycles, adjust induction programmes, or revise observation frameworks. Improvement activity is tracked through indicators such as staff retention, learner progress measures, lesson observation outcomes, and consistency of curriculum delivery.

AT LEVELS 6-8

(practice development, innovation and experimentation), the provider may develop new workforce models. This could include digital CPD platforms, micro-credential approaches, peer-led scholarship models, or cross-provider mentoring networks. Early experimentation may test feasibility and uptake, refining delivery before larger-scale implementation.

AT LEVELS 9-10

(pilots, trials and evaluation), the approach is tested across multiple sites, departments or learner cohorts, supported by structured evaluation. At this stage, the question shifts from 'does this seem promising?' to 'does it work reliably, for whom, and under what implementation conditions?'. Mixed-method evaluation may be used to understand outcomes and mechanisms, ensuring that professional development is linked to demonstrable learner benefit rather than participation alone.

AT LEVEL 11

(practitioner research and action research), professional learning becomes explicitly knowledge-generating. Practitioner researchers may publish case studies on coaching models, peer observation, or subject-specific pedagogy improvement, with transparent methodologies and ethical reflection. Action research is particularly powerful here, as it integrates intervention design, implementation, reflection and evidence-building into iterative cycles (Reason and Bradbury, 2008).

AT LEVEL 12

(applied research), evidence can be aggregated across providers or regions to build more robust conclusions about effective professional development models in FE and skills. This might include comparative evaluation of CPD approaches, analysis of workforce pipelines, or modelling the relationship between staff development and learner progression outcomes.

AT LEVEL 13

(academic research), deeper theory-building studies may examine professional identity, effective CPD design, or the relationship between workforce development and learner outcomes, strengthening the conceptual infrastructure that informs applied and practitioner inquiry.

AT LEVEL 14

(evidence synthesis and knowledge mobilisation), findings across multiple studies are consolidated through systematic review or structured evidence synthesis, producing sector-level guidance. At this point, professional learning becomes codified into professional standards, CPD frameworks, leadership programmes, and inspection-facing quality improvement models. This is the stage at which professional development is no longer dependent on individual organisational capacity but becomes part of the sector's shared infrastructure.

This example illustrates how everyday practice can travel through the taxonomy: beginning with local reflective practice, becoming structured inquiry and evaluated innovation, and ultimately generating transferable, codified evidence that strengthens professionalism across the system.

REFERENCES

- Argyris, C. and Schön, D.A. (1978) *Organizational learning: A theory of action perspective*. Reading, MA: Addison-Wesley.
- Ball, S.J. (2015) Education, governance and the tyranny of numbers. *Journal of Education Policy*, 30(3), pp. 299–317.
- Bathmaker, A.-M. (2013) 'Researching the FE sector: Policy, practice and critical perspectives', *Journal of Vocational Education & Training*, 65(2), pp. 145–153.
- Biesta, G. (2010) *Good education in an age of measurement: Ethics, politics, democracy*. Boulder, CO: Paradigm Publishers.
- Cohen, L., Manion, L. and Morrison, K. (2018) *Research methods in education*. 8th edn. London: Routledge.
- Eraut, M. (1994) *Developing professional knowledge and competence*. London: Falmer Press.
- Eraut, M. (2004) 'Informal learning in the workplace', *Studies in Continuing Education*, 26(2), pp. 247–273.
- Freire, P. (1970) *Pedagogy of the Oppressed*. New York: Continuum.
- Flyvbjerg, B. (2001) *Making social science matter: Why social inquiry fails and how it can succeed again*. Cambridge: Cambridge University Press.
- Flyvbjerg, B. (2006) 'Five misunderstandings about case-study research', *Qualitative Inquiry*, 12(2), pp. 219–245.
- Greany, T. and Higham, R. (2018) *Hierarchy, markets and networks: Analysing the 'self-improving school-led system' agenda in England*. London: UCL IOE Press.
- Hargreaves, A. and Fullan, M. (2012) *Professional capital: Transforming teaching in every school*. New York: Teachers College Press.
- Smith V, Husband, G. and Smith, M. (2026) Why research in further education and skills matters. *Education +Training* 68 (3): 297–302.
- Keep, E. (2018) 'The English skills system: Confusions, contradictions and chaos', *Local Economy*, 33(2), pp. 141–157.
- Kemmis, S., McTaggart, R. and Nixon, R. (2014) *The action research planner: Doing critical participatory action research*. Singapore: Springer.
- Lewin, K. (1946) 'Action research and minority problems', *Journal of Social Issues*, 2(4), pp. 34–46.
- Lincoln, Y.S. and Guba, E.G (1985) *Naturalistic inquiry*. Newbury Park, CA: Sage.
- Nonaka, I. and Takeuchi, H. (1995) *The knowledge-creating company: How Japanese companies create the dynamics of innovation*. Oxford: Oxford University Press.
- Ofsted (2026) Ofsted Areas of Research Interest. Available at: <https://www.gov.uk/government/publications/ofsted-areas-of-research-interest> (Accessed: 13/04/26).

- Patton, M.Q. (2008) *Utilization-Focused Evaluation*. 4th edn. Thousand Oaks, CA: Sage.
- Pawson, R. and Tilley, N. (1997) *Realistic evaluation*. London: Sage.
- Petticrew, M. and Roberts, H. (2006) *Systematic reviews in the social sciences: A practical guide*. Oxford: Blackwell.
- Polanyi, M. (1966) *The tacit dimension*. London: Routledge & Kegan Paul.
- Reason, P. and Bradbury, H. (eds.) (2008) *The SAGE Handbook of Action Research: Participative Inquiry and Practice*. 2nd edn. London: Sage.
- Rowley, J. (2007) 'The wisdom hierarchy: Representations of the DIKW hierarchy', *Journal of Information Science*, 33(2), pp. 163–180.
- Sachs, J. (2016) 'Teacher professionalism: Why are we still talking about it?', *Teachers and Teaching*, 22(4), pp. 413–425.
- Schön, D.A. (1983) *The reflective practitioner: How professionals think in action*. New York: Basic Books.
- Senge, P.M. (2006) *The fifth discipline: The art and practice of the learning organization*. 2nd edn. London: Random House Business.
- Stenhouse, L. (1975) *An introduction to curriculum research and development*. London: Heinemann.
- Wenger, E. (1998) *Communities of practice: Learning, meaning, and identity*. Cambridge: Cambridge University Press.



Dr Vikki Smith,
Chief Professionalism Officer,
Education Training Foundation (ETF)

Dr Mark Hyde,
Interim Director of Professional Statuses and Research,
Education Training Foundation (ETF)



As the professional body for the FE and skills workforce, ETF sets standards, provides continuing development opportunities, and fosters an inclusive community of educators, leaders and support teams.

Our work is supported through funding by the Department for Education (DfE), provider funding and professional membership.

157-197 Buckingham Palace Road, London SW1W 9SP
020 3740 8280
enquiries@etfoundation.co.uk
ETFOUNDATION.CO.UK

Company registration number (England and Wales): 08540597. Charity number: 1153859