February 2026

The Evidence Imperative

Knowing Every Learner as the Foundation of Adaptive Education

A White Paper on **Explore Learning's** Evidence-Based Approach to AI and Technology Adoption in Childhood Education.

"The only good learning is that which is in advance of development."

— Lev Vygotsky, in Cole et al. (Eds.), Mind in Society (1978)

Dr Hisham Ihshaish
Head of Data and AI

With contributions from:
Lewis Cherry
Dr Selvan Gnanakumaran
Rosie Hamps
Charlotte Gater
Emily Rathge
Richard Reynolds

research@explorelearning.co.uk

Contents

Executive Summary

The global education technology market is projected to reach USD 348 billion by 2030, with the UK sector alone expected to grow to USD 31.15 billion (Grand View Research, 2025). This expansion, driven significantly by advances in artificial intelligence, machine learning, and generative AI technologies such as large language models (LLMs), brings both promise and peril.

Recent meta-analyses reveal that a range of AI-based interventions demonstrates significant positive effects on academic performance (effect size d = 0.92) (Dong et al., 2025), yet simultaneously raises concerns about higher-order cognition: research links AI over-reliance to reduced critical thinking depth (Stadler et al., 2024) and cognitive fixation that undermines creative confidence (Habib et al., 2024).

This paradox, improvement in measurable outcomes alongside potential erosion of higher-order cognitive capacities, demands a fundamentally different approach to AI in education. One that prioritises evidence over hype, knowing the learner over algorithmic convenience, and human guidance over technological displacement.

Core Thesis

True personalisation is impossible without rigorous, ongoing assessment of each child's zone of proximal development, learning rate, and mastery trajectory.

Technology serves this understanding; it does not replace it.

This white paper articulates Explore Learning's vision and methodology: an evidence-based framework where continuous learner understanding is the prerequisite for adaptive learning. We are critical of educational technology (EdTech) solutions that promise personalisation without the evidential infrastructure to deliver it.

Understanding every learner at every step is not a feature. It is the foundation.

The UK Education Context: Challenges and Opportunities [1]

1.1 A system under pressure

The UK education system faces significant structural challenges that technology alone cannot solve but may help address. Understanding this context is essential for responsible EdTech development.

The disadvantage gap remains substantial.

The Education Policy Institute's 2025 annual report shows the GCSE-level attainment gap between disadvantaged pupils[2] and their peers narrowed marginally from 19.2 months in 2023 to 19.1 months in 2024, but remains substantially wider than pre-pandemic levels (Education Policy Institute, 2025). At primary level, disadvantaged pupils are 10 months behind their peers by the end of Key Stage 2 (age 10-11). In 2024, just 43% of disadvantaged pupils achieved a standard pass (grade 4 or above) in both English and maths GCSE, compared to 73% of non-disadvantaged pupils. When measured against a strong pass (grade 5 or above) – the government's benchmark for 'good' attainment – the gap is starker still: just 26% of disadvantaged pupils, compared to 53% of their peers (DfE, 2024b). These gaps emerge early and widen as children progress through education, affecting future employment and lifetime earnings.

The government has recognised this challenge directly. In a January 2026 policy announcement, the Department for Education indicated that up to 450,000 children from disadvantaged backgrounds could benefit from AI-powered tutoring tools, noting that "access to tutoring is deeply unequal, with children from wealthier families far more likely to benefit" (DfE, 2026a).

Children with SEND face compounded disadvantage.

The Sutton Trust's October 2025 research reveals a troubling pattern of "double disadvantage": children eligible for free school meals are significantly overrepresented within the SEND cohort, and within that cohort, outcomes diverge sharply by socio-economic background (Sutton Trust, 2025). In 2023/24, only 7.5% of children with an EHCP who were eligible for FSM secured grade 4 or above in English and Maths, compared to 17.3% of those with an EHCP but not eligible for FSM. This intersection of disadvantage and special educational needs represents one of the most pressing challenges in UK education, and one where evidence-based technology may offer meaningful support (explored further in Section 6.2).

Teacher recruitment and retention remain in crisis.

The National Foundation for Educational Research reports that 90% of teachers considering leaving cite high workload as a factor (NFER, 2025). The teacher vacancy rate is six times higher than pre-pandemic levels. For all but one of the past ten years, the Department for Education has missed its target for secondary school teacher training recruitment. The government's pledge to recruit 6,500 additional teachers faces significant uncertainty (NAO, 2025).

[1] Education policy is devolved: England, Scotland, Wales, and Northern Ireland each have distinct curricula, qualification frameworks, and AI in education policies. The data and policy discussion in this section relate primarily to England. Many underlying challenges - attainment gaps, teacher workforce pressures, the need for evidence-based technology - are shared across the home nations, though specifics differ. In Scotland, for example, teacher workforce dynamics diverge from the English picture, with shortages concentrated in specific subjects alongside insufficient permanent posts for newly qualified teachers (BBC Scotland, 2026). For comparative attainment data across all four nations, see the Joseph Rowntree Foundation's education monitoring reports.

[2] In UK education policy, 'disadvantaged pupils' refers to children eligible for free school meals (FSM), children looked after by the local authority, or children previously in care. The disadvantage gap is measured against non-disadvantaged peers.

Digital skills gaps persist despite curriculum reform.

Computing became a mandatory national curriculum subject in 2014, yet the Curriculum and Assessment Review (2025) found that 23% of UK businesses report a lack of basic digital competence among employees, with 37% lacking required advanced digital skills (DfE, 2025). Time spent teaching computing in secondary schools has declined since the curriculum change, and girls make up only 21% of the GCSE Computer Science cohort (Computing at School, 2024).

These challenges create both the imperative and the opportunity for evidence-based educational technology that supports learning rather than merely automating assessment. The following section examines the broader EdTech landscape, before we turn to the theoretical foundations and practical approaches that guide Explore Learning's response.

1.2 UK policy on AI in education

The UK government has taken an increasingly active role in shaping AI's integration into education, with significant policy developments in early 2026.

The Department for Education's guidance on generative AI acknowledges both potential and risk:

> *"If used safely, effectively and with the right infrastructure in place, AI can support every child and young person, regardless of their background, to achieve at school and college and develop the knowledge and skills they need for life"* (DfE, 2024).

Investment in evidence-based development.

In January 2026, the Education Secretary announced a £23 million investment to expand the EdTech Testbeds pilot into a four-year programme, recruiting over 1,000 schools and colleges to test AI and EdTech tools in classroom settings (DfE, 2026b). The explicit aim is to build "genuine evidence about what's working, the cream of education tech and AI rising to the top." This represents a significant policy commitment to evidence-based adoption rather than technology for its own sake.

New safety standards for generative AI.

The DfE's updated Generative AI Product Safety Standards (Jan 2026) establish requirements spanning content safety, cognitive development, emotional safeguarding, monitoring, data protection, and more (DfE, 2026c). Products must use progressive disclosure rather than providing final answers by default, a regulatory recognition that learning requires productive struggle, not frictionless delivery. They must track and report when learners offload thinking to the system, and must not anthropomorphise AI or foster emotional dependence. The standards also require detection of learner distress with appropriate safeguarding pathways, and prohibit manipulative strategies including sycophancy and dark patterns. Taken together, these represent a maturing regulatory understanding that educational AI requires design principles fundamentally distinct from general-purpose applications.

Focus on disadvantage and inclusion.

The government's AI tutoring initiative explicitly targets equity: "The government is running a tender for industry to co-create AI tutoring tools with teachers, with the goal of bringing these tools to a similar level of quality [to human tutoring], so that we can offer, at scale, the kind of personalised one-to-one support often only available to a privileged few" (DfE, 2026a). This framing, technology as a means of democratising access to high-quality support, aligns with the evidence on tutoring effectiveness while acknowledging current inequalities.

SEND provision.

The government published its Schools White Paper, Every Child Achieving and Thriving, in February 2026. The paper outlines proposed reforms to the SEND system, including changes intended to strengthen early support in mainstream settings and improve consistency of provision. These proposals are subject to consultation and phased implementation, meaning their impact on outcomes for pupils with SEND will take time to unfold.

Meanwhile, a dedicated SEN Identification and Support Research and Innovation Challenge has been launched under the government's £500 million R&D Missions Accelerator Programme to explore data-driven tools for earlier identification of needs (GOV.UK, 2025).

The Curriculum and Assessment Review recommends that AI literacy, data science, and the ethical use of technology be embedded across the reformed curriculum, with a new Computing GCSE replacing the current Computer Science qualification to better reflect the breadth of digital skills needed (DfE, 2025).

The EdTech Landscape: Promise and Peril

The UK context outlined above creates the conditions for educational technology's growth. But growth alone does not guarantee benefit. This section examines the market dynamics driving EdTech expansion and, fundamentally, what the evidence shows about effectiveness.

2.1 Market expansion and technological innovation

The education technology sector is experiencing substantial growth, driven significantly by advances in artificial intelligence. According to Grand View Research, the global EdTech market was valued at USD 163.49 billion in 2024 and is projected to reach USD 348.41 billion by 2030, representing a compound annual growth rate of 13.3% (Grand View Research, 2025). Primary and secondary education[3] dominates with approximately 39% of market share. In the United Kingdom specifically, the EdTech market is expected to reach USD 31.15 billion by 2030, with primary and secondary education representing 43.76% of sector revenue (Grand View Research UK, 2025).

The emergence of generative AI and large language models (LLMs), including tools like ChatGPT, Claude, and Gemini, has accelerated innovation. These technologies enable new capabilities: automated lesson planning, personalised feedback at scale, adaptive content generation, and conversational tutoring. The DfE notes that *"generative AI has demonstrated that it can help the education workforce by reducing some of the administrative burdens that hard-working teachers, staff and school leaders face"* (DfE, 2024).

However, this capital influx creates powerful dynamics. Venture capital expectations typically demand 5–7 year returns, creating pressure for features demonstrating immediate, measurable impact, typically standardised test improvements, over supporting complex, long-term developmental outcomes (Selwyn, 2019). The result is what Muller (2018) terms "metric fixation": the reduction of education to easily quantifiable outcomes.

2.2 What the evidence shows

The most comprehensive recent meta-analysis examining AI's effect on academic achievement, published in Computers and Education: Artificial Intelligence (Dong et al., 2025), synthesised 29 empirical studies comprising 2,657 participants.

These studies examined AI applications focused on student personalised learning pathways, adaptive tutoring systems, intelligent learning environments, and personalised content delivery, rather than AI tools primarily designed to enhance teacher productivity.

The findings reveal a significant positive effect on academic performance, with an overall effect size of $d = 0.924$, a large effect by conventional standards (effect size measures and thresholds are defined in Table 1). A separate meta-analysis focusing specifically on generative AI (Ma et al., 2025), synthesising 34 studies, found a combined effect size of $g = 0.68$, with cognitive outcomes at $g = 0.795$. Notably, the effect size for purpose-built AI systems across the broader literature ($d = 0.92$) exceeds that observed for generative AI specifically ($g = 0.68$). This pattern is consistent with the OECD's distinction between purpose-built educational tools and general-purpose AI (OECD, 2026), and reinforces the case for pedagogically grounded design over generic deployment.

[3] In the UK, this covers Reception through Year 13 (ages 4–18), equivalent to the US term 'K–12' (Kindergarten through 12th grade).

A Stanford-led meta-analysis of 119 studies on educational technology interventions for early literacy[4] provides further granularity (Silverman et al., 2025). Table 1 synthesises the evidence across these sources.

Table 1: Effect sizes of AI and educational technology by domain and intervention type.

Domain	Effect size[a]		Source
AI on academic performance (broad)	$d = 0.92$	▲▲	Dong et al. (2025), 29 studies
GenAI on learning outcomes	$g = 0.68$	▲	Ma et al. (2025), 34 studies
Writing proficiency (ages 5–10[b])	$g = 0.81$	▲▲	Silverman et al. (2025), 119 studies
Decoding skills (ages 5–10[b])	$g = 0.33$	△	Silverman et al. (2025)
Reading comprehension (ages 5–10[b])	$g = 0.23$	△	Silverman et al. (2025)
Creative thinking	Negative[c]	▼	Habib et al. (2024)
Critical thinking (over-reliance)	Negative[c]	▼	Stadler et al. (2024)

▲▲ Large (≥ 0.80) ▲ Medium (0.50–0.79) △ Small (0.20–0.49) ▼ Negative

[a] Effect sizes are reported as Cohen's d or Hedges' g, both standardised measures of the magnitude of an intervention's impact. By convention: ≥ 0.80 is large, 0.50–0.79 is medium, and 0.20–0.49 is small (Cohen, 1988). [b] UK equivalent: Reception to Year 5 (primary school). [c] Qualitative; no standardised effect size reported.

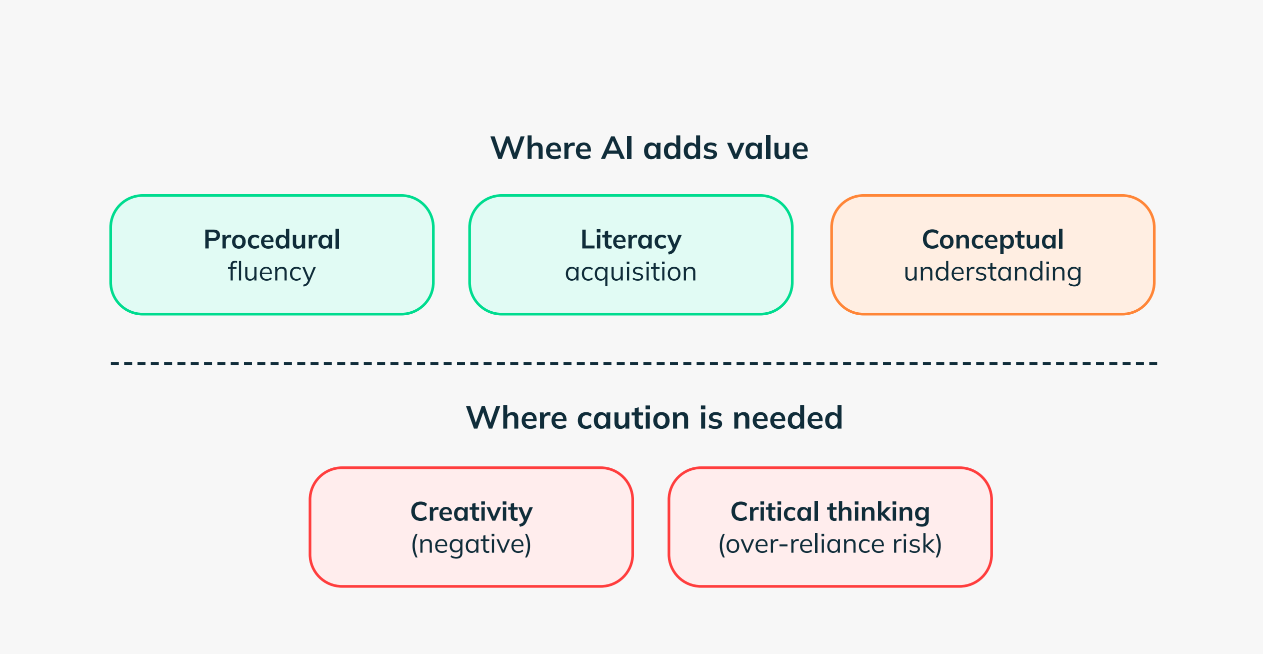
The pattern across these meta-analyses is consistent: AI and educational technology demonstrate strong effects for procedural, rule-based skills and foundational literacy but diminishing, or negative, effects as cognitive complexity increases. In a randomised experiment with 91 university students, Stadler et al. (2024) found that those using ChatGPT-3.5 for scientific inquiry reported significantly lower cognitive load than those using traditional search engines, yet produced weaker reasoning and argumentation in their conclusions, cognitive ease, at a direct cost to analytical depth.

Habib et al. (2024), studying undergraduates using AI for divergent thinking tasks in a creative problem-solving course, found that while AI improved fluency in idea generation, it also produced cognitive fixation, participants reported difficulty generating their own ideas after seeing AI outputs, and reduced creative confidence.

These studies examined university students with fully developed metacognitive capacities; for younger learners whose executive function and self-regulation are still developing, the risks of cognitive dependency are arguably greater. Jose et al. describe this as the "cognitive paradox" of AI in education: AI functions simultaneously as cognitive amplifier and inhibitor, enhancing performance at lower taxonomic levels while risking erosion of the higher-order thinking it cannot model (Jose et al., 2025).

The heterogeneity of effects is itself an important finding. Wu et al. (2025) found that effect sizes varied dramatically by subject domain within the same meta-analysis, from negligible to large, underscoring that blanket claims about 'AI in education' are meaningless. Evidence must be tool-specific and context-specific.

[4] The Silverman et al. (2025) meta-analysis covers quasi-experimental and experimental studies published between 2010 and 2023, evaluating educational technology interventions on literacy outcomes for children aged 5–10 (UK Reception through Year 5).



Sources: Dong et al. (2025), Ma et al. (2025), Silverman et al. (2025), Habib et al. (2024), Stadler et al. (2024)

Figure 1: AI effectiveness varies substantially by cognitive domain. Strong effects for procedural skills; limited or negative effects for higher-order thinking.

2.3 Interpreting the evidence

The debate around educational technology's effectiveness will, and should, continue. Healthy scepticism protects against hype cycles and vendor overclaims. However, this debate must be grounded in accurate representation of evidence, and in recognition that 'educational technology' encompasses vastly different interventions with different outcomes.

The evidence demands nuance. The Silverman et al. meta-analysis found positive effects across literacy outcomes, from g = 0.23 for reading comprehension to g = 0.81 for writing proficiency, with the lead author concluding: "There isn't a single answer to whether digital technologies support literacy. The question is: *which products, with which characteristics, under which conditions?*" Furthermore, for students from low socioeconomic backgrounds, certain programmes showed larger positive effects, a finding with significant equity implications that reinforces the case for evidence-based, rather than blanket, adoption.

Popular coverage of educational technology often conflates distinct intervention types, treats correlation as causation, and misrepresents nuanced research findings. The Silverman meta-analysis, for instance, was characterised in prominent media commentary as showing that educational technology delivers only marginal benefits, a claim that directly contradicts the authors' actual conclusion that investing in educational technology to support literacy is warranted, with effectiveness varying by product characteristics, target skills, and student populations. This gap between primary research and its public interpretation illustrates precisely why rigorous engagement with the evidence base matters.

The OECD's Digital Education Outlook 2026 reinforces this nuance (OECD, 2026). The report identifies the "mirage of false mastery": students using general-purpose AI saw short-term gains, but performance dropped when AI access was removed.

The conclusion is not that AI has no place in education, but that general-purpose AI carries risks when used without pedagogical purpose. Purpose-built educational tools, grounded in established learning theory, show more promising results. The report calls for *"a shift away from use of off-the-shelf chatbots in favour of purpose-built educational GenAI systems, designed in conjunction with teachers."*

This evidence landscape points toward clear principles: effectiveness depends on design, implementation, and pedagogical grounding. The question is not whether educational technology 'works' but which technologies, designed how, deployed under what conditions, support learning for which children. The theoretical frameworks outlined in the following section provide the foundation for answering this question.

Theoretical Foundations: From Theory to Practice

The evidence reviewed above, the importance of appropriate challenge, scaffolding that fades as competence develops, the limits of technology in higher-order cognition, and the irreplaceable role of human relationships, echoes foundational learning theories developed decades before AI entered classrooms. We ground our development in these established frameworks not because they are fashionable, but because the evidence consistently validates them.

3.1 Vygotsky's Zone of Proximal Development

Lev Vygotsky's concept of the zone of proximal development (ZPD), articulated in the posthumously compiled Mind in Society (Vygotsky, 1978), provides the theoretical cornerstone for understanding optimal learning conditions. Vygotsky defined the ZPD as:

> *"The distance between the actual developmental level as determined by independent problem solving and the level of potential development as determined through problem solving under adult guidance, or in collaboration with more capable peers" (Vygotsky, 1978, p. 86).*

This definition contains profound implications for educational technology:

First, learning is not passive absorption but active construction occurring through interaction. The ZPD explicitly requires a "more knowledgeable other", whether human or, potentially, algorithmic, to mediate between current capability and potential.

Second, the zone is dynamic, not static. As the learner develops, the zone shifts. What was once in the ZPD moves to independent capability, and new challenges enter the zone. This dynamism demands continuous assessment, not one-time placement.

Third, instruction that falls below or above the ZPD is suboptimal. Content that is too easy fails to promote cognitive growth; content that is too difficult leads to frustration and disengagement.

3.2 Scaffolding: Structured support for learning

The term "scaffolding", introduced by Wood, Bruner, and Ross (1976), describes temporary support structures that enable learners to operate within their ZPD (Wood et al., 1976). Effective scaffolding requires knowing precisely where the learner stands: what they can do independently, what they can do with support, and what remains beyond reach. This is where technology can offer distinctive value: not only the capacity to track performance patterns across time, but the ability to deliver personalised scaffolding at scale, adjusting difficulty, pacing, and support responsively and continuously for each learner.

Such scaffolding must, by design, fade as competence develops; the ultimate goal is always independent mastery, not perpetual dependency on support.

Crucially, effective scaffolding does not mean eliminating difficulty. Learning requires productive struggle, the experience of encountering challenge, making errors, and working through them. Mistakes are not failures; they are essential learning signals that reveal misconceptions and prompt the deeper processing that consolidates understanding. This is why the DfE's Generative AI Product Safety Standards require that AI tools use progressive disclosure rather than providing immediate answers (DfE, 2026c): children need space to think, to attempt, and to discover for themselves. Technology that short-circuits this process, however efficient it appears, undermines the very learning it claims to support.

3.3 Bloom's Taxonomy: A framework for cognitive mastery

The taxonomy of educational objectives, originally developed by Bloom et al. (1956) and subsequently revised by Anderson & Krathwohl (2001), provides a hierarchical framework for understanding cognitive skills. The revised taxonomy identifies six cognitive processes, remember, understand, apply, analyse, evaluate, and create, progressing from foundational to higher-order thinking.

Mapping these levels to educational practice clarifies what is at stake. At the foundational levels, remembering involves retrieving facts and definitions; understanding requires explaining concepts in one's own words; applying means using knowledge in new situations. These are the cognitive tasks at which AI demonstrably excels, pattern recognition, spaced repetition, procedural guidance, and immediate corrective feedback.

The higher levels demand qualitatively different cognition: analysing requires decomposing arguments, identifying assumptions, and distinguishing evidence from opinion; evaluating involves making judgements under uncertainty, weighing competing values, and critiquing reasoning; creating demands original synthesis, combining ideas in novel ways to produce something that did not previously exist.

This hierarchy maps directly onto the empirical evidence. The large positive effects reported by Dong et al. (2025) and Ma et al. (2025) predominantly reflect gains at the foundational levels, precisely where AI's capacity for adaptive practice and personalised scaffolding creates genuine value. The negative effects reported by Habib et al. (2024) on creative thinking and Stadler et al. (2024) on analytical depth correspond to the higher taxonomic levels where current AI systems lack the contextual understanding, metacognitive awareness, and genuine comprehension that human cognition provides. The pattern is not incidental; it reflects a fundamental asymmetry between what AI can and cannot do.

The implication for EdTech is clear: AI can effectively support mastery of foundational cognitive skills, remembering, understanding, applying, freeing human educators to focus on higher-order development, analysing, evaluating, creating, where their expertise is irreplaceable.

3.4 Social Constructivism: Learning as social process

Vygotsky's broader social constructivist framework emphasises that cognitive development is fundamentally social. Children develop higher mental functions through interaction with more capable others, gradually internalising what was once external dialogue.

This has direct implications for EdTech design: **digital scaffolding can support but not replace human mediation.** While technology can provide certain forms of scaffolding, adaptive difficulty, immediate feedback, personalised pacing, it cannot provide the full range of social, emotional, and motivational support that human interaction offers. The most effective educational technology recognises this limitation and positions itself as augmenting, not replacing, human relationships.

The Primacy of Knowing the Learner

The theoretical frameworks above converge on a single practical requirement: to scaffold learning within the ZPD, to position tasks at the right level of Bloom's hierarchy, and to provide the social mediation that constructivism demands, a system must first know the learner. Without rigorous, continuous assessment of where each child stands, these frameworks remain abstractions. With it, they become the architecture for adaptive education.

The Evidence Imperative

You cannot adapt to a learner you do not know.

Many EdTech products claim "personalised learning" without the evidential infrastructure to deliver it. True personalisation requires knowing, with precision and currency, each learner's current capability, learning rate, areas of strength and weakness, and optimal challenge threshold.

Without this knowledge, "adaptation" becomes guesswork, adjusting difficulty based on crude metrics like percentage correct, without understanding ***why*** a student succeeded or struggled.

Any system claiming to deliver personalised learning must be able to identify each learner's ZPD accurately and continuously. This requires knowing:

- The child's current capability level across different domains
- Their learning rate, how quickly they acquire and retain new skills
- The scaffolding they need to progress
- How these factors change over time

4.1 Learning within the ZPD

We operationalise this theoretical framework through systematic measurement of performance on cognitive tasks within the learner's estimated zone of proximal development. When a child is positioned appropriately within their ZPD, they can complete tasks successfully with appropriate scaffolding, not so easy that no learning occurs, not so difficult that frustration sets in.

Performance is tracked over time across multiple sessions, providing an evolving picture of how each child responds to tasks at varying difficulty levels. This longitudinal view, rather than any single snapshot, reveals the child's true learning trajectory.

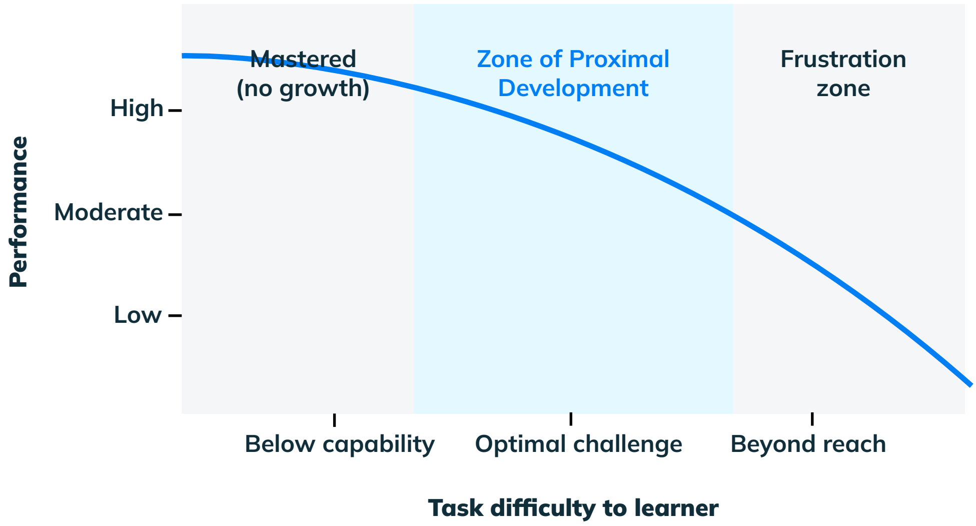


Figure 2: Performance on cognitive tasks as a function of task difficulty. In the optimal ZPD (blue), tasks are challenging enough to promote learning while remaining achievable with appropriate support. Grey zones represent ineffective positioning, either too easy, high performance but no cognitive growth, or too hard, declining performance and disengagement.

The ZPD is not static, it shifts as the child develops. What was challenging last month may now fall within independent capability; what seemed impossible may now be within reach with scaffolding. This dynamic nature is precisely why continuous monitoring matters: a child assessed once and placed accordingly will inevitably drift from optimal positioning as they grow.

4.2 Aptitude as a dynamic measure

A critical insight from our work is that aptitude is not a fixed trait but a dynamic characteristic that changes over time. Traditional approaches often assess children once and assume static capability. This fundamentally misunderstands learning.

Within any given year group, performance on cognitive tasks varies across a range, some children progress quickly, some more slowly, and most fall in a broad middle. We stratify learners into aptitude bands based on observed performance over time, **but these bands are not fixed labels.** They are dynamic classifications that update based on ongoing data. A child's band membership is recalculated at regular intervals, providing confidence in assignments while allowing for movement as children develop.

This dynamic tracking is central to our approach. Systems that capture how aptitude changes over time, responding to intervention, scaffolding, and development, can deliver adaptive learning. Systems that treat aptitude as fixed cannot.

The practical implications are significant:

- Children who improve are recognised and appropriately challenged
- Children who struggle receive additional scaffolding before falling behind
- No child is left in an inappropriate learning zone due to outdated assessment
- Progress monitoring becomes a continuous process, not a periodic event

4.3 Why this matters

Consider two children, both showing similar performance on initial assessment, and subsequently placed at the same cognitive level. Without additional data, any system would treat them identically. But the initial performance alone conceals fundamentally different learning profiles: one is a rapid learner for whom this assessment sat comfortably within existing capability; the other is working at the boundary of their current reach. Their demonstrated performance is identical; their learning trajectories, and therefore their optimal pathways, are not.

Generic systems would give both children identical content, identical pacing, identical scaffolding. One would be under-challenged; the other, potentially overwhelmed.

An evidence-based approach differentiates from the first interaction. By tracking not just what children get correct but how they get there, response time, error patterns, consistency across sessions, we build a dynamic model of each learner that enables genuine personalisation.

Explore Learning's Approach

Our Framework:
Evidence-Based Adaptive Learning

Explore Learning's approach is built on a foundational conviction: personalisation requires genuine understanding of each learner. We have developed an integrated evidence-driven architecture that operationalises the theoretical frameworks of Vygotsky, Bloom, and social constructivism into practical, measurable systems. Six guiding principles shape how we implement this conviction.

5.1 Our guiding principles

Continuous assessment, not episodic.
A child assessed once will inevitably drift from optimal positioning as they develop. Assessment must be ongoing, capturing change as it happens. This ensures learning pathways respond to who the child is now, not who they were at enrolment.

Dynamic aptitude, not fixed.
Two children at the same curriculum position may require entirely different scaffolding, and their needs will change over time. Dynamic aptitude tracking ensures children who accelerate are appropriately challenged, while those who need support receive it before falling behind.

AI supports, humans guide.
The synthesis of human and artificial intelligence is more powerful than either alone. Technology delivers measurable progress in skill acquisition; human tutors develop the creativity, critical thinking, and resilience that AI cannot.

Parents as partners.
When families understand and can support their child's learning, outcomes improve. Engaged, informed parents become active participants in their child's progress, extending learning beyond the session.

Evidence before expansion.
We develop and deploy technology only when we have evidence it helps children learn. Every capability we deploy has evidence behind it; every limitation is acknowledged.

Retention, not just acquisition.
Knowledge that cannot be retrieved independently weeks or months later has not been truly learned. Our systems track not only initial skill acquisition but subsequent retention and transfer.

The cumulative result: every child is understood, appropriately positioned, and supported, at every step of their learning journey.

5.2 The Four EdTech Pillars

Our technological development rests on four interconnected pillars that operationalise these principles, enabled by a foundation of meaningful parent engagement:

1	2	3	4
Initial Assessment	**Aptitude-Led Mastery**	**Progress Monitoring**	**Human-AI Synthesis**
Fast, accurate placement	Dynamic ZPD positioning	Real-time adaptation	Tech amplifies, humans guide

Enabled by parent / carer engagement

Informed • Empowered • Participating

Knowing every learner at every step

Pillar 1: Initial Assessment

Effective personalisation begins at enrolment. Our initial assessment is designed to be:

- **Fast:** Completed within a single session, minimising burden on child and family
- **Accurate:** Calibrated against national curriculum standards and validated through longitudinal outcome data
- **Comprehensive:** Mapping not just current level but highlighting relative strengths, weaknesses, and areas requiring focused attention
- **Actionable:** Generating immediate placement recommendations and initial learning pathway

The assessment employs an adaptive testing methodology, enabling efficient identification of the child's capability boundary.

Pillar 2: Aptitude-Led Mastery

Rather than merely advancing children to higher skill levels as they learn, we aim to enhance their mastery potential at each level through personalisation and scaffolding calibrated to their dynamic aptitude profile.

The result is **cohesive learning groups:** within each aptitude band, children learn at similar rates with aligned mastery criteria. During periodic checks, students whose aptitude has shifted, whether improving or requiring additional support, move to the appropriate band. This ensures every child receives scaffolding matched to their current needs, not their historical profile.

Pillar 3: Continuous Progress Monitoring

Initial assessment provides a starting point; continuous monitoring ensures the pathway remains optimal:

- **Session-by-session performance:** Every interaction generates data on accuracy and consistency across cognitive tasks
- **Learning rate evolution:** Aptitude band membership is recalculated regularly, capturing developmental change
- **Skill mastery progression:** Movement through Bloom's cognitive hierarchy is tracked
- **Engagement indicators:** Patterns suggesting disengagement trigger human review
- **Predictive insight:** For exam-stage learners, performance predictions based on current trajectory enable targeted interventions before, not after, difficulties become entrenched

Our development trajectory extends this monitoring to incorporate additional signals, including error pattern analysis and response characteristics, that will enable even more responsive adaptation. The principle remains constant: technology generates insight and delivers personalised scaffolding at scale; humans interpret, guide, and develop the child beyond what algorithms can reach.

Pillar 4: Human-AI Synthesis

The question is not whether to use AI but how to integrate it purposefully. Our model distributes responsibilities clearly:

AI handles	Humans provide
Adaptive difficulty calibration	Motivational coaching
Personalised scaffolding at scale	Progress interpretation
Pattern detection and prediction	Conceptual explanation
Immediate feedback on practice	Emotional support
Routine skill drilling	Relationship building

Shared: assessment design, intervention decisions, mastery verification

Figure 3: Distribution of responsibilities. AI amplifies human capability; it does not replace it.

This represents purposeful technology integration, deploying AI at the core of the educational process, where evidence demonstrates it creates the most value, rather than bolted onto the periphery for novelty or marketing differentiation. Where evidence shows strong effects, academic performance, writing proficiency, foundational literacy, technology can and should deliver personalised scaffolding, adaptive challenge, and immediate feedback at a scale no human team alone could sustain. Where evidence shows risk, higher-order cognition, creative development, emotional growth, human expertise must lead.

Consider what happens in practice: a tutor reads hesitation in a child's posture, recognises the difference between productive struggle and genuine distress, and adjusts not just the task but the emotional register of the interaction. No algorithm can interpret these non-verbal cues with the sensitivity that effective teaching demands.

The outcome of this synthesis:

- **Technology delivers:** personalised scaffolding, adaptive challenge calibration, predictive insights, and progress monitoring, the evidence-based foundations of effective skill acquisition
- **Humans develop:** the capabilities that matter most in an AI-rich world, creativity, critical thinking, resilience, communication, and the confidence to learn

Neither alone achieves this. Technology without human guidance risks the dependency and cognitive erosion the evidence warns against. Human expertise without technological support cannot scale to meet every child where they are. The combination is what makes adaptive education effective.

5.3 Parent/Carer Engagement: The Enabling Foundation

Research consistently demonstrates that parental engagement significantly impacts educational outcomes (Education Endowment Foundation, 2024). Yet the current system creates an opacity problem: parents or carers often have limited information on how their child is doing, and even less clarity on what they can do to support learning at home. This leaves many parents feeling lost or stressed, a challenge that is even more acute for families navigating SEND.

This is an under-discussed part of the education system, and Explore Learning is focused on addressing it. With appropriately designed AI-powered tools, the opportunity can be profound: **informing, empowering, and enabling** parents in ways not previously possible at scale.

Our Parent/Carer App and coaching tools bridge this gap:

- **Real-time progress visibility:** Parents see learning trajectories, skill development, and areas of focus, not raw metrics, but interpretable summaries that make sense
- **Personalised recommendations:** AI coaching tools suggest specific activities parents can support at home, targeted to current learning goals and calibrated to the child's ZPD
- **Transparent communication:** Parents understand not just what their child is learning, but why, connecting daily activities to longer-term progress
- **For exam-stage learners:** Personalised predictions based on current performance and trajectory, with specific recommendations for achieving target grades

The AI coaching tools inform and guide, they complement human support rather than replacing it. Tutors remain central to nuanced communication: concern flags, strategy discussions, complex questions. But routine updates, practice reminders, and milestone celebrations can be handled at scale, ensuring every parent has visibility into their child's learning journey.

The result: when children are learning within their ZPD, with parents engaged and informed, the evidence suggests higher engagement and more progress. This is not technology for technology's sake, it is technology in service of the human relationships that drive learning.

The architecture described above represents where we are today. But the evidence base is evolving, the policy landscape is maturing, and the needs of learners are not static. The following section outlines where we are heading, and the principles that will guide us there.

Future Directions: Digital Skills and Inclusive Support

6.1 Digital skills: An essential foundation

The UK government's Curriculum and Assessment Review has emphasised that *"all young people should be equipped with the digital capabilities required for an increasingly technology- and AI-enabled future"* (DfE, 2025). Digital skills are integral to the UK economy, yet 8.5 million adults lack basic digital competence (Good Things Foundation, 2024). Evidence-based educational technology has a role to play in building these foundational capabilities, but only if it develops understanding rather than surface-level tool familiarity.

6.2 SEND: Technology in service of inclusion

As outlined in Section 1.2, children with special educational needs and disabilities face compounded disadvantage, particularly those from lower-income backgrounds. The scale of this challenge warrants emphasis: nationally, 25.7% of children are eligible for free school meals (FSM), yet this figure rises to 39.3% for those receiving SEND support and 43.8% for those with Education, Health and Care Plans (EHCPs). Research by Impetus found that children from disadvantaged backgrounds with SEND who have low qualifications are 180% more likely to be not in education, employment or training (NEET) than average (Sutton Trust, 2025). This is not inevitable, but addressing it requires earlier identification and more responsive support than current systems typically provide.

The challenge and opportunity

The research evidence for technology-supported identification is encouraging: machine learning models can achieve high accuracy in detecting conditions such as dyslexia and dyscalculia through analysis of performance patterns (Zaibi et al., 2024; Sedmidubsky et al., 2025; Fink, 2025), and the OECD's 2025 working paper provides a comprehensive analysis of both the potential and limitations of such approaches (OECD, 2025). The UK government has recognised this potential, announcing a dedicated SEN Identification and Support Research and Innovation Challenge, delivered by UKRI in partnership with the DfE as part of the broader £500 million R&D Missions Accelerator Programme, to explore how data-driven tools could support earlier identification of needs (GOV.UK, 2025).

For children from disadvantaged backgrounds who face additional barriers to formal assessment, this matters most. Continuous monitoring could be transformative, identifying patterns that warrant further investigation, rather than waiting for children to fall far enough behind to trigger referral.

Our approach

We see significant potential for evidence-based systems to support SEND, through early signal detection, adaptive scaffolding, and pattern recognition across populations. Explore Learning's integrated approach, spanning assessment, learning provision, progress monitoring, evidence collection, and support for children, parents, and schools, positions us to explore this potential systematically.

Measuring the impact of SEND interventions is notoriously difficult, yet increasingly demanded by schools and funders. Our continuous monitoring architecture provides the longitudinal data that accountability requires. However, we are clear-eyed about limitations: technology can flag patterns warranting investigation; it cannot diagnose. Technology can provide adaptive scaffolding; it cannot replace specialist human expertise.

6.3 The limits of technology: Why human expertise remains essential

Throughout this paper, we have emphasised what technology can do. It is equally important to be clear about what it cannot.

Context requires human judgement.
Algorithmic outputs, whether flagging potential SEND, recommending interventions, or predicting outcomes, must be interpreted by professionals who understand the child's full context: family circumstances, educational history, social and emotional factors. No algorithm captures the whole child.

Relationships cannot be automated.
The motivational, emotional, and social dimensions of learning remain fundamentally human. A child who feels known, supported, and believed in will learn differently from one who feels like a data point. This is why tutors remain central to what we do.

Ethical considerations demand human oversight.
Labelling children algorithmically raises significant concerns. The UK government's initiative specifically emphasises that any tools developed will be "*subject to rigorous data protection, safeguarding protocols, and ethical approvals*" (GOV.UK, 2025). We endorse this caution.

No algorithm is perfect.
False positives and negatives are inevitable. Human professional judgement remains essential for accurate identification and appropriate response.

The goal is not to replace human expertise but to augment it, ensuring no child struggles invisibly when earlier support could help, while preserving the human relationships that make education meaningful.

What we reject

7.1 Technology for technology's sake

We are critical of EdTech that deploys AI, including generative AI and large language models, primarily for marketing advantage without clear evidence of educational benefit. The question is never *"does it use AI?"* but *"does it help children learn?"*

7.2 Displacement narratives

We reject the claim that AI will or should replace human educators. The evidence is unambiguous: AI's strengths and human educators' strengths are complementary, not substitutive. The DfE is clear: *"AI could never be a substitute for teachers' professional judgement and the personal relationships they have with their students"* (DfE, 2024).

7.3 Black-box algorithms

Educational decisions should be explicable. When we place a child in an aptitude band or recommend an intervention, tutors and parents should understand why. The DfE's Safety Standards require that products track and report cognitive offloading and use progressive disclosure rather than opaque answer provision (DfE, 2026c), principles that demand algorithmic transparency. We reject opaque decision-making that cannot be interrogated or explained.

7.4 Superficial "personalisation"

We are sceptical of platforms claiming personalisation while delivering essentially identical content to all learners, adjusted only by surface difficulty. The OECD identifies a *"mirage of false mastery"* in which students using general-purpose AI show short-term gains that vanish when AI access is removed (OECD, 2026). True personalisation requires the deep, continuous learner understanding that this paper describes.

Conclusion: Innovation with responsibility

Our Commitment

Explore Learning is committed to the **innovation of education, scientifically and responsibly.** We believe technology can transform learning outcomes. We also believe this transformation must be grounded in evidence, guided by established learning theory, and integrated with irreplaceable human expertise.

The evidence reviewed in this paper reveals a paradox at the heart of AI in education. Technology demonstrates strong positive effects on foundational skills, yet simultaneously risks eroding the higher-order cognition, creative confidence, and analytical depth that education exists to develop. Resolving this paradox requires not less technology, but better technology: systems designed with pedagogical purpose, grounded in continuous learner understanding, and integrated with human expertise where it is irreplaceable.

This is not an abstract concern. With a disadvantage gap of 19.1 months at GCSE level (Education Policy Institute, 2025) and compounded barriers for children with SEND (Sutton Trust, 2025), the stakes are concrete. If access to evidence-based, personalised learning technology remains unequal, AI risks widening the very gaps it could help close. The government's commitment to reaching 450,000 disadvantaged pupils (DfE, 2026a) reflects this urgency, but realising that commitment demands the kind of rigorous, evidence-based approach this paper describes.

8.1 The convergence of evidence and policy

The policy landscape is moving decisively toward the approach this paper describes, and compliance with these emerging frameworks matters, both for the integrity of educational technology and for the children it serves.

The DfE's £23 million EdTech Testbeds expansion (Jan 2026) explicitly prioritises "genuine evidence about what's working" through rigorous classroom evaluation (DfE, 2026b). The updated Generative AI Product Safety Standards (Jan 2026) establish requirements that align directly with evidence-based principles: progressive disclosure rather than answer provision, cognitive offloading tracking, safeguards against emotional dependency, and protection from manipulative engagement strategies (DfE, 2026c). The government's commitment to deploying AI tutoring tools for up to 450,000 disadvantaged pupils (DfE, 2026a), co-created with teachers, reflects a maturing understanding that technology in education must be purposeful, equitable, and safe. International research, including the OECD's call for "purpose-built educational GenAI systems, designed in conjunction with teachers" (OECD, 2026), reinforces this direction.

This convergence validates the approach Explore Learning has been developing. We did not design our systems to comply with these emerging standards; we designed them because the evidence pointed this way. That our architecture, continuous assessment, dynamic aptitude tracking, human-AI synthesis, transparent reporting, meets and in many cases anticipates these requirements is not incidental. It reflects a shared understanding, across policy, research, and practice, of what responsible educational technology requires.

8.2 Where we are heading

The architecture described in this paper is not an end state. We see it as the foundation for an evolving ecosystem in which richer evidence, error pattern analysis, response characteristics, retention trajectories, enables increasingly precise scaffolding; predictive insight supports earlier intervention for those who need it most; continuous monitoring helps identify children with special educational needs before they fall far enough behind to trigger referral; and parents become genuine partners in their child's learning rather than passive recipients of reports. Throughout, the principle remains constant: technology generates insight and delivers personalised support at scale; humans interpret, guide, and develop the child beyond what algorithms can reach.

8.3 An invitation

The global education technology market will reach USD 348 billion by 2030. Generative AI will continue to advance. The question is not whether technology will be part of education's future but what kind of technology, deployed how, in service of what goals.

Our answer: evidence-based technology, deployed in synthesis with human expertise, in service of genuinely understanding and supporting every learner. Technology that helps children master foundational skills while humans develop the creativity, critical thinking, resilience, and communication that will define success in an AI-rich world.

We invite schools, educators, researchers, policymakers, and innovators who share this commitment to engage with us, not as observers, but as partners in shaping the future of education every child deserves.

This is the evidence imperative.

This is how we lead in the age of AI, not by chasing hype, but by building understanding.

Knowing every learner. At every step.

Explore Learning Ltd.
research@explorelearning.co.uk

References

Department for Education. (2024). Key stage 4 performance, Academic year 2023/24 (revised). Explore Education Statistics. https://explore-education-statistics.service.gov.uk/find-statistics/key-stage-4-performance/2023-24

Anderson, L. W., & Krathwohl, D. R. (Eds.). (2001). A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives. New York: Longman.

BBC Scotland. (2026). Teachers on supply list for years or leaving country due to job shortages. BBC News, 19 January 2026.

Muller, J. Z. (2018). The Tyranny of Metrics. Princeton, NJ: Princeton University Press.

Bloom, B. S., Engelhart, M. D., Furst, E. J., Hill, W. H., & Krathwohl, D. R. (1956). Taxonomy of Educational Objectives: The Classification of Educational Goals. Handbook I: Cognitive Domain. New York: David McKay Company.

Cohen, J. (1988). Statistical Power Analysis for the Behavioral Sciences (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Computing at School. (2024). Decoding the decade: The time for reform in computing education is now. https://www.computingatschool.org.uk

Department for Education. (2024). Generative artificial intelligence (AI) in education. GOV.UK. https://www.gov.uk/government/publications/generative-artificial-intelligence-in-education

Department for Education. (2025). Curriculum and Assessment Review: Final report and government response. GOV.UK.

Department for Education. (2026). 450,000 disadvantaged pupils could benefit from AI tutoring tools. GOV.UK, 26 January 2026. https://www.gov.uk/government/news/450000-disadvantaged-pupils-could-benefit-from-ai-tutoring-tools

Department for Education. (2026). Education Secretary speech at Bett UK Conference. GOV.UK, 21 January 2026. https://www.gov.uk/government/speeches/education-secretary-speech-at-bett-uk-conference

Department for Education. (2026). Generative AI: product safety standards. GOV.UK, 19 January 2026. https://www.gov.uk/government/publications/generative-ai-product-safety-standards

Dong, L., Tang, X., & Wang, X. (2025). Examining the effect of artificial intelligence in relation to students' academic achievement: A meta-analysis. Computers and Education: Artificial Intelligence, 8, 100400. https://doi.org/10.1016/j.caeai.2025.100400

Education Endowment Foundation. (2024). Working with parents to support children's learning: Guidance report. London: EEF. https://educationendowmentfoundation.org.uk/education-evidence/guidance-reports/supporting-parents

Education Policy Institute. (2025). Annual Report 2025: Disadvantage. https://epi.org.uk

Fink, N. (2025). Enhancing the early detection of learning disabilities through machine learning: Achieving over 91% accuracy. SSRN Working Paper. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5134981

Good Things Foundation. (2024). Basic digital skills: Expert overview. https://www.goodthingsfoundation.org

GOV.UK. (2025, November 20). Children with SEND to benefit from earlier intervention as government backs research into tools to identify special educational needs sooner. Department for Science, Innovation and Technology.

Grand View Research. (2025). Education Technology Market Size, Share & Trends Analysis Report 2025–2030. https://www.grandviewresearch.com/industry-analysis/education-technology-market

Grand View Research. (2025). UK Education Technology Market Size & Outlook, 2030. https://www.grandviewresearch.com/horizon/outlook/education-technology-market/uk

Habib, S., et al. (2024). How does generative artificial intelligence impact student creativity? Journal of Creativity, 34(1), 100072. https://doi.org/10.1016/j.yjoc.2023.100072

Jose, B., Cherian, J., Verghis, A. M., Varghise, S. M., & Joseph, S. (2025). The cognitive paradox of AI in education: Between enhancement and erosion. Frontiers in Psychology, 16, 1550621. https://doi.org/10.3389/fpsyg.2025.1550621

Ma, S., Lei, H., & Fang, W. (2025). The effects of generative artificial intelligence on student learning outcomes: A meta-analysis. Journal of Computer Assisted Learning. Advance online publication. https://doi.org/10.1111/jcal.70117

National Audit Office. (2025). Government's 6,500-teacher pledge faces uncertainties as student numbers surge. NAO Press Release, 29 April 2025.

National Foundation for Educational Research. (2025). Teacher Labour Market in England Annual Report 2025. https://www.nfer.ac.uk

OECD. (2025). Leveraging artificial intelligence to support students with special education needs. OECD Working Paper.

OECD. (2026). OECD Digital Education Outlook 2026: Exploring Effective Uses of Generative AI in Education. OECD Publishing, Paris. https://doi.org/10.1787/062a7394-en

Selwyn, N. (2019). Should robots replace teachers? AI and the future of education. Cambridge: Polity Press.

Sedmidubsky, J., Dostalova, N., Svaricek, R., & Culemann, W. (2025). INSIGHT: Combining fixation visualisations and residual neural networks for dyslexia classification from eye-tracking data. Dyslexia, 31(1), e1801.

Silverman, R. D., Keane, K., Darling-Hammond, E., & Khanna, S. (2025). The effects of educational technology interventions on literacy in elementary school: A meta-analysis. Review of Educational Research, 95, 972–1012. https://doi.org/10.3102/00346543241261073

Stadler, M., Bannert, M., & Sailer, M. (2024). Cognitive ease at a cost: LLMs reduce mental effort but compromise depth in student scientific inquiry. Computers in Human Behavior, 160, 108386. https://doi.org/10.1016/j.chb.2024.108386

O'Regan, C., & Latham, K. (2025). Double disadvantage? Socio-economic inequalities in the SEND system. London: The Sutton Trust. https://www.suttontrust.com/our-research/double-disadvantage/

Vygotsky, L. S. (1978). Mind in Society: The Development of Higher Psychological Processes (M. Cole, V. John-Steiner, S. Scribner, & E. Souberman, Eds.). Cambridge, MA: Harvard University Press.

Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. Journal of Child Psychology and Psychiatry, 17(2), 89–100.

Wu, T., Wang, C., & Li, B. (2025). Effects of AI in blended learning: A meta-analysis. Frontiers in Psychology, 16, 1504636.

Zaibi, A., et al. (2024). Early detection of learning disabilities through handwriting analysis and machine learning. Procedia Computer Science, 246, 2149–2158.