Interaction Design and Architecture(s) Journal - IxD&A, N.67, 2025, pp. 10 - 23
DOI: 10.55612/s-5002-067-001

The Death of Cognitive Conflict? AI, Learning, and the Case for Disruptive Pedagogy

Ines Di Loreto[1,2]

[1] Laboratory "Computer Science and Digital Society" (LIST3N), team "Technologies et Pratiques" (TePra) Troyes University of Technology, 12 rue Marie Curie, CS 42060, 10004 Troyes
[2] European Culture and Technology Laboratory (ECT Lab+), European University of Technology, European Union

ines.di_loreto@utt.fr

Abstract. AI-powered educational technologies promise to enhance personalization, efficiency, and accessibility. However, this paper argues that modern AI-driven learning environments increasingly prioritize seamless automation over cognitive conflict, uncertainty, and disruption, key mechanisms for learning. Drawing on theories of cognitive disequilibrium, prediction error, and systems thinking, we examine how early AI learning systems embraced productive struggle, while contemporary adaptive platforms focus on minimizing errors and optimizing user experience. Finally, we explore the role of Human-Computer Interaction (HCI) in shaping AI's educational future, advocating for intentional disruptions that foster engagement and conceptual change.

Keywords: Artificial Intelligence in Education (AIEd), Learning Theories, Human-Computer Interaction, Disruptive Learning, AI and Pedagogy.

1 Introduction

The use of Artificial Intelligence in education has gained significant interest in recent years, with public discourse highlighting its potential to enhance learning experiences, improve accessibility, and provide personalized instruction at scale [1, 2]. On the scientific front, AI-powered educational technologies (including Intelligent Tutoring Systems (ITSs), adaptive learning platforms, and generative AI tools) have long promised to revolutionize teaching and learning by providing data-driven insights, automating feedback, and tailoring content to individual learners' needs (e.g., [3,4,5]). Yet, despite over 30 years of research in Artificial Intelligence in Education (AIEd), many educators remain unsure how to leverage these technologies pedagogically or understand their meaningful impact on teaching and learning [6]. Despite the growing interest, research in this field remains largely dominated by descriptive and pilot studies with a technological focus, as well as quantitative, quasi-experimental methods. The limited presence of longitudinal studies suggests that there is still significant opportunity to explore new approaches to better understand AI's potential [6, 7].

Finally, as AI increasingly mediates the learning process, concerns have emerged regarding its pedagogical assumptions and their alignment with established theories of learning [6, 8, 9].

Building on these concerns, this paper examines the role of cognitive conflict, uncertainty, and disruption in learning, particularly in the context of AI-driven education. Section 2 explores how theories of learning emphasize cognitive disequilibrium as a driver of conceptual change, tracing its roots from Piaget's work to contemporary computational models. Section 3 examines early AI-based learning systems, such as Intelligent Tutoring Systems (ITSs) and expert systems, highlighting how they integrated cognitive conflict and error-driven learning. In contrast, Section 4 analyzes modern adaptive AI platforms, which largely prioritize efficiency and personalization over productive struggle. Section 5 extends the discussion to broader implications within AI and Human-Computer Interaction (HCI), questioning whether AI's trajectory toward automation and frictionless interaction risks undermining deeper learning processes. Finally, Section 6 offers concluding reflections on the challenges and opportunities of designing AI-driven education that embraces, rather than eliminates, cognitive friction.

2 The Role of Uncertainty and Conflict in Learning: A Historical Perspective

Understanding how individuals process and integrate conflicting information has long been a central question in cognitive and educational psychology. Theories of learning have actually consistently emphasized the role of uncertainty, contradiction, and disruption in driving cognitive development and adaptation.

This section traces the evolution of these theories, highlighting how learning is not a passive accumulation of knowledge, but an active, iterative process driven by cognitive conflict, uncertainty, and environmental disruptions.

2.1 The Origins: Cognitive Disequilibrium and Conceptual Change

We will begin our historical overview with Jean Piaget's theory of *cognitive disequilibrium*. This theory postulates that learning is not a passive accumulation of knowledge, but an active process driven by the conflict between new experiences and pre-existing mental schemas. Piaget asserted that when individuals encounter information that contradicts their prior knowledge, they enter a state of cognitive disequilibrium, a psychological discomfort that prompts them to revise, reorganize, or expand their understanding [10]. This process occurs through assimilation, where new information is incorporated into existing cognitive structures, and accommodation, where mental models are adjusted to resolve inconsistencies. This disequilibrium, that we could also define as a cognitive imbalance [11] is thus a necessary catalyst for intellectual growth, as learners are motivated to restore equilibrium by modifying their knowledge framework to incorporate the unexpected information [12]. Piaget's work laid the foundation for constructivist learning theories and the idea that knowledge is actively constructed by learners through active engagement with novel and sometimes contradictory information, rather than passively absorbed.

In the same years Leon Festinger's [13] *cognitive dissonance* theory explored the psychological discomfort experienced when individuals encounter information inconsistent with their beliefs, leading them to adjust their cognitions to reduce this discomfort. It's worth to note that Festinger explores how people deal with conflicting beliefs and attitudes, particularly in social and decision-making contexts and not in learning contexts. However, while Festinger's work is more focused on the social psychology of belief and behavior change, his theory helps explain why individuals might resist new information or selectively reinterpret it.

Another psychologist who contributed to this topic in the 1950/60s was Daniel Berlyne [14]. Berlyne affirmed that *curiosity* serves as a motivational prerequisite for exploratory behavior [15]. He distinguished between different types of curiosity, notably perceptual curiosity, which is awakened by novel stimuli and diminishes with exposure, and epistemic curiosity, a desire for knowledge that motivates learning and intellectual exploration. Stimulus properties such as novelty, complexity, uncertainty, and conflict affect an individual's motivation to explore. These variables can create a sense of uncertainty or conflict, prompting exploratory behavior to resolve the ambiguity. However, if the discrepancy between expectations and actual information is too large, expectation change decreases [16].

Back to learning theories and expanding on Piaget's idea in more recent years, Chinn and Brewer [17] detailed how learners react to anomalous data (and thus to information that directly contradicts their existing beliefs). According to the authors, learners exhibit a range of responses when confronted with unexpected findings, including outright rejection, reinterpretation, or complete restructuring of their conceptual framework [17]. Their findings help to explain, for example, why students who hold naive physics misconceptions about motion may initially resist experimental evidence contradicting their beliefs, and why, with sufficient cognitive engagement and scaffolding, they may undergo conceptual change, leading to a deeper and more accurate understanding.

Less widely known is D'Mello et al. [18] work which explored the effects of *confusion* as a learning mechanism. Their research demonstrated that when learners encounter uncertainty, contradictions, or unexpected outcomes, they are more likely to engage in deep cognitive processing, provided that the confusion is effectively resolved. The study found that well-designed learning environments that strategically introduce cognitive dissonance can lead to stronger conceptual retention and more adaptive problem-solving skills. This supports the idea that productive struggle, rather than immediate clarity, fosters durable learning.

To summarize, in these theories learning is not a passive accumulation of knowledge, but a dynamic process shaped by tension between existing beliefs and new information. When individuals encounter contradictions, they experience a cognitive disequilibrium that compels them to revise and expand their understanding. Rather than impeding learning, uncertainty and contradiction fuel intellectual growth, much like curiosity draws individuals toward exploration and discovery. The struggle to reconcile contradictions is thus not merely an obstacle but a necessary force in learning. When uncertainty is carefully introduced and effectively resolved, it fosters engagement, strengthens retention, and enhances problem-solving skills. Yet, responses to such challenges are not uniform: some may reject conflicting information outright, while others reinterpret or fully restructure their understanding, revealing the difficulty of conceptual change.

Learning, then, is not best supported by immediate clarity, but by the productive discomfort that drives individuals to question, adapt, and ultimately construct more robust knowledge frameworks.

2.2 From Cognitive Conflict to Prediction Errors: A Computational Perspective on Learning

While cognitive disequilibrium and conceptual change theories emphasize the role of psychological discomfort and conceptual restructuring in learning, recent advances in cognitive science have formalized these processes through mathematical and neuroscientific frameworks. The mathematical turn in cognitive science (including for example predictive processing, Bayesian inference, and error-driven learning) revives and refines early computational models of the brain. However, while early theories treated the brain as a calculator, today's models propose a more dynamic,

Interaction Design and Architecture(s) Journal - IxD&A, N.67, 2025, pp. 10 - 23
DOI: 10.55612/s-5002-067-001

probabilistic system that constantly adjusts its internal representations based on real-world interactions, emphasizing neural computation and embodied cognition. In Bayesian learning models, the brain does not treat knowledge as fixed or absolute, but rather as a set of evolving probabilities that change with new experiences. Instead of storing rigid facts, the brain assigns degrees of certainty to different beliefs, adjusting them based on incoming information. This means that learning is a continuous process of refining expectations, rather than simply replacing old knowledge with new facts and that learners do not passively absorb information but actively construct and refine mental models through continuous prediction and correction.

For example, according to Clark [19], the human brain operates as a predictive engine, continuously generating expectations about incoming sensory information based on prior experiences. When these expectations are violated, a prediction error occurs, signaling that the learner's internal model of the world requires updating [19]. This mechanism aligns with Bayesian inference principles, where new information is weighted against prior knowledge, and the brain updates its predictions accordingly.

Other studies in cognitive neuroscience suggest that these Bayesian principles are fundamental to human reasoning, motor learning, and perceptual adaptation (e.g.; [20, 21]. For example, research in physics education [22, 23, 24] has shown that when students predict the outcome of an experiment and then observe a result that contradicts their expectation, they are more likely to reflect, question their assumptions, and revise their conceptual models than if they were merely provided with an answer upfront.

Interesting enough, these Bayesians models are reflected in neuroscientific studies. Schultz et al. [25], for example, demonstrated that when individuals encounter an unexpected outcome, the dopamine system in the brain is activated, reinforcing learning and prompting behavioral adjustments [25]. In particular, their work on reinforcement learning suggests that unexpected rewards produce stronger neural encoding than anticipated ones, reinforcing the importance of designing educational experiences that introduce controlled elements of surprise. This neuroscientific evidence underscores that unexpected learning moments are biologically wired to enhance retention and adaptability [26, 27], reinforcing the need for pedagogical strategies that introduce productive uncertainty.

In summary, prediction errors and computational models expand on earlier theories of learning by showing that rather than simply resolving contradictions, learners are engaged in an ongoing probabilistic refinement of knowledge, where unexpected events serve as catalysts for deeper reflection and restructuring of mental models.

2.3 Learning in Complex Systems: Adapting to Feedback, Disruptions, and Uncertainty

While Bayesian models and predictive processing provide a computational framework for individual learning, real-world learning often takes place in dynamic, interconnected systems, where surprises and disruptions expose hidden causal relationships. In these environments, learning is not only about refining internal predictions but also about adapting to complex, unpredictable external forces. Unlike the individual updating of beliefs described in 2.2, systems learning requires responding to interdependent variables, feedback loops, and emergent behaviors that defy straightforward cause-and-effect reasoning. Unexpected challenges in such systems force learners to rethink causal assumptions, refine mental models, and develop more sophisticated problem-solving strategies [28, 29].

Systems Thinking, introduced by Meadows [30], offers a framework for understanding how disruptions reveal hidden relationships in complex systems. The theory emphasizes that real-world problems are interconnected and often behave in non-linear ways, thus recognizing that small changes can create disproportionate ripple effects, requiring learners to shift from linear to *systemic* reasoning. When learners engage with dynamic, complex systems (such as ecosystems, economic models, or, we could add, AI decision-making processes) unexpected disruptions can force them to reassess cause-and-effect relationships and develop more sophisticated mental models [31].

One of the most well-known applications of systems-based learning is climate science education. Research has shown that students struggle to grasp climate feedback loops and emergent behaviors unless they are presented with interactive simulations where variables behave unpredictably [29]. This reinforces Meadows' assertion that true understanding does not arise from memorizing isolated facts but from grasping the intricate, often chaotic interactions between elements in a system.

To summarize, theories of cognitive disequilibrium (2.1) and prediction errors (2.2) have demonstrated how learning thrives on contradiction and uncertainty at an individual level. However, real-world learning rarely happens in isolation: it unfolds in complex, interconnected systems, where feedback loops, emergent behaviors, and non-linear relationships challenge traditional cause-and-effect reasoning. Rather than refining internal predictions alone, learners navigating complex systems must continuously reassess their assumptions in response to external perturbations.

3 Early AI in Learning: Expert Systems and Intelligent Tutoring Systems

As we have seen in Section 2, theories of learning have long emphasized the role of cognitive conflict, uncertainty, and environmental disruptions in shaping knowledge acquisition. These foundational theories have not only influenced educational psychology but have also informed the development of AI-driven learning environments from the 1970s to the 1990s.

Early AI applications in education, in fact, sought to replicate expert reasoning and adapt instructional content to individual learners, incorporating principles of cognitive conflict and error-driven learning.

The following sections examine how expert systems and Intelligent Tutoring Systems (ITSs) emerged as early attempts to bring these pedagogical theories into AI-powered learning environments.

3.1 Expert Systems and Early AI in Education

Expert systems were among the first AI applications in education, designed to simulate human expert decision-making within specific domains [32]. These systems relied on rule-based reasoning and symbolic AI to guide learners through problem-solving processes by structuring knowledge representation and inference mechanisms [33, 34]. One of the earliest and most influential expert systems was MYCIN [34], a medical diagnostic system that provided explanations for its reasoning. Although MYCIN was designed for medical diagnosis, it influenced intelligent tutoring systems by demonstrating how AI could model expert reasoning and provide explanations.

Building on the success of MYCIN, in fact, Clancey [35], developed GUIDON, which adapted MYCIN's rule-based structure for instructional purposes. GUIDON allowed learners to interact with the system, receive expert feedback, and refine their problem-solving strategies. This approach aligned with constructivist learning theories, particularly in its emphasis on active engagement with expert knowledge and structured problem-solving. Rather than relying on passive information transfer, GUIDON encouraged learners to explore domain knowledge interactively and develop reasoning skills through guided explanations. However, these systems were not designed to introduce contradictions. Instead, they helped users navigate expert rules and receive explanations, more in line with scaffolding and guided instruction rather than cognitive conflict.

3.2 The Rise of Intelligent Tutoring Systems

Parallel to expert systems, the 1970s and 1980s saw the development of Intelligent Tutoring Systems, which aimed to provide individualized, adaptive instruction based on a learner's cognitive state [36]. Unlike traditional computer-assisted instruction, ITSs sought to dynamically adjust content delivery and problem-solving feedback, making them more responsive to learners' needs [37].

One of the earliest ITSs, SCHOLAR [38], used semantic networks to structure knowledge representation in geography education. SCHOLAR allowed learners to engage in natural language dialogue, making it one of the first AI systems to incorporate interactive question-answering as a pedagogical tool. However, this system did not explicitly introduce contradictions or challenge misconceptions as it was more of a structured knowledge retrieval system.

More relevant to the purposes of this paper, SOPHIE [39] introduced learning through error detection, in which students' incorrect solutions were analyzed to generate targeted feedback. This method reflected research on conceptual change and the role of cognitive conflict in learning [17].

Finally, a particularly influential ITS was *Andes*, developed in the 1990s to teach physics [40]. *Andes* implemented constraint-based tutoring, ensuring that errors led to adaptive, problem-specific feedback rather than generic responses. Andes was one of the most advanced tutoring systems developed in the late 1990s and early 2000s, and it contributed significantly to research on cognitive models of learning, adaptive feedback, and student problem-solving behavior [40]. Its design was consistent with cognitive theories suggesting that learning occurs most effectively when students experience productive struggle and have opportunities to resolve conceptual contradictions.

3.3 Cognitive Conflict and AI-Driven Learning in Early Systems

Many early AI learning systems were directly influenced by constructivist learning theories (e.g.; Piaget) and cognitive disequilibrium models, emphasizing that contradictions between prior knowledge and new information drive conceptual change. Some AI-based tutoring systems employed *Socratic dialogue techniques*, encouraging students to articulate reasoning and reflect on inconsistencies in their understanding [41].

Socratic dialogue is an instructional method rooted in the questioning approach used by Socrates, where an instructor (or system) poses a series of carefully structured questions to guide learners toward deeper reflection and self-correction. Instead of providing direct answers, these systems prompted students to justify their reasoning, identify gaps in their knowledge, and refine their mental models through critical thinking. For example, *AutoTutor* [42] used natural language processing to engage students in Socratic questioning, challenging misconceptions and prompting self-explanation. By fostering metacognitive awareness and conceptual change, these systems leveraged cognitive conflict as a mechanism for deeper learning.

One significant development in this regard was the "*BUGGY model*" approach [43], which simulated common student misconceptions to highlight cognitive conflicts. By intentionally exposing learners to conflicts between their intuitive understanding and correct principles, these systems fostered conceptual change, aligning with Chinn and Brewer's [17] research on responses to anomalous data.

Additionally, AI-driven simulation-based learning environments were developed to integrate controlled disruptions. These systems allowed learners to interact with dynamic, unpredictable scenarios, reinforcing adaptability and real-time problem-solving. For example, early AI-driven training simulations, such as those used in procedural learning and virtual environments, introduced unexpected variables to challenge students' assumptions and encourage flexible thinking [44, 45].

3.4 Transition from Rule-Based AI to Machine Learning (1990s)

By the 1990s, AI in education began shifting from rule-based expert systems to machine learning-driven models, incorporating probabilistic reasoning and adaptive learning. The already cited *AutoTutor* [42] was one of the first ITSs to integrate natural language processing, allowing more fluid interactions between students and AI tutors.

This shift also aligned with the rise of Bayesian inference models, which framed learning as a probabilistic process of updating beliefs based on new information [46]. Rather than relying on fixed rules, AI systems began using predictive algorithms to estimate learner comprehension and provide tailored instructional scaffolding [3]. We will delve further into this matter in the next section.

To summarize, the historical evolution of AI in learning reveals a trajectory from rule-based expert systems to adaptive, machine learning-driven educational platforms. Early AI applications used cognitive conflict, conceptual change, and productive struggle as central elements for effective learning. These systems, while constrained by their reliance on symbolic AI, pioneered individualized, disruption-based learning experiences.

It's worth to note that these models were recognized in academic research settings, but they were not widely applied or known to the general public. Their adoption remained limited by computational constraints, lack of widespread computing resources, and the challenge of integrating AI into traditional education at the time.

4 Learning Through Unexpected Outcomes in Technology Enhanced Learning – Recent years.

The theories explored in Section 2 converge on a central insight: unexpected outcomes are a fundamental driver of learning and learning is not a passive accumulation of knowledge but an active, disruption-driven process. The presented theories emphasize that effective learning environments should intentionally introduce controlled discrepancies, cognitive dissonance, and unpredictable feedback loops to engage learners in deeper, more adaptive cognitive processing. As we saw in Section 3, the same theories have historically been applied to AI-based learning. But what about more recent years?

This section will focus on the use of disruptions to enhance cognitive engagement and foster adaptive learning strategies in Technology Enhanced Learning (TEL) after 2010. In the second part of this section, we will address the role of disruption in AI-powered learning systems in recent days.

4.1 Unexpected Outcomes in Technology-Enhanced Learning

As far as the author of this paper knows, in recent years only a small subset of research in Technology-Enhanced Learning has incorporated controlled disruptions as a means of fostering deeper engagement and adaptive learning. However, the findings of these studies are highly interesting. These unexpected elements act as cognitive perturbations, prompting learners to revise mental models, develop problem-solving skills, and engage in metacognitive reflection [47, 48].

Gamification, in particular, has emerged as a widely used strategy within TEL to enhance motivation and learning retention, often through dynamic goal-setting, feedback variability, and the strategic introduction of obstacles [49]. While gamified systems frequently emphasize predictable reinforcement mechanisms (e.g., immediate rewards and achievements), research suggests that incorporating unexpected feedback and variable reinforcement can lead to deeper engagement and resilience. For example, Sailer and Homner [50] conducted a meta-analysis on gamification in education and found that challenge-based learning elements, including unpredictable game mechanics and controlled failures, encouraged cognitive reevaluation and persistence. However, they also noted that excessive unpredictability could lead to frustration and disengagement, highlighting the need to carefully calibrate uncertainty in learning environments. Similarly, Seaborn and Fels [51] explored how dynamic goal structures and unexpected setbacks in gamification improve problem-solving skills by forcing learners to adapt to changing conditions rather than relying on memorization strategies.

Beyond gamification, simulated learning environments provide a controlled yet unpredictable setting in which learners must navigate uncertainty and adapt in real time. For example, studies in medical education and emergency response training illustrate how introducing unexpected patient complications or crisis scenarios enhances diagnostic accuracy and decision-making [52]. Similarly, research in aviation training demonstrates that flight simulators incorporating unpredictable system failures and environmental hazards significantly improve pilots' crisis management skills [53].

These findings reinforce the argument that deliberate unpredictability in simulations strengthens transfer learning, improving learners' ability to apply knowledge in novel real-world contexts [54].

4.2 AI Driven Systems

AI in education shows in recent years even less examples of the usage of disruption strategies. As far as we know, current AI-based learning platforms prioritize, in fact, efficiency, personalization, and scaffolding, often avoiding cognitive conflict or failure-based learning. Most AI-driven commercial educational platforms (such as Khan Academy, Duolingo, Coursera, and so on) leverage personalized learning paths that optimize instruction to match a learner's current level of understanding. These systems analyze user performance and dynamically adjust content to ensure maximum efficiency in skill acquisition. While this approach increases engagement and accessibility, it largely avoids the kind of cognitive conflict necessary for conceptual change.

For example, Duolingo's adaptive algorithm aims to keep learners in a state of optimal *engagement* by providing incremental difficulty progression and immediate corrective feedback, reducing frustration but also minimizing conceptual struggle. While this may help users acquire vocabulary efficiently, it does not actively encourage deep linguistic reflection or metacognitive engagement with language structure. Similarly, Khan Academy's mastery-based progression model ensures students advance only when they demonstrate proficiency in a given skill but does not inherently challenge them with conflicting information or force them to wrestle with conceptual contradictions.

On the scientific end, however, some research has shown promising results.

AI-powered adaptive learning systems can dynamically adjust content delivery based on learners' interactions while also incorporating unpredictability to stimulate engagement and cognitive flexibility. Two meta-analyses by Holmes et al. [2] and

VanLehn [3] review studies on AI-driven educational technologies, demonstrating how AI-based environments can introduce controlled discrepancies between expected and actual learning outcomes, prompting deeper conceptual revision and metacognitive processing. Rooted in reinforcement learning principles, some AI tutors strategically delay feedback, provide counterintuitive hints, or present information in unexpected sequences to challenge learners' assumptions and encourage adaptive reasoning [55].

Interesting enough, generative AI tools, such as large language models (LLMs) and intelligent tutoring systems, can *unintentionally* amplify learning unpredictability by generating variable and sometimes inconsistent responses to identical queries, exposing learners to multiple perspectives and problem-solving strategies [56, 57].

In particular in [57] the authors "played" with large LLMs probabilistic answers. In LLMs, in fact, the same input can produce different outputs, sometimes leading to inconsistencies. While often seen as a drawback, the study found that exploiting such inconsistencies, though reducing users' perceived trust in AI, actually improved their comprehension of the information provided. These findings suggest that instead of hiding inconsistencies, AI systems could present them transparently to encourage critical thinking and more informed engagement with AI-generated content.

However, the opacity of black-box AI models presents a challenge to learners navigating AI-driven environments. Many adaptive learning algorithms operate as complex neural networks, making their decision-making processes difficult to interpret [58]. The unpredictability of AI-generated reasoning requires students to engage in deeper critical analysis to understand the logic behind recommendations and feedback. While this process can enhance higher-order thinking skills, excessive unpredictability without adequate transparency may lead to confusion or disengagement [59]. Addressing this issue requires improving the interpretability of AI systems in education, ensuring that learners can trace and understand AI-driven instructional decisions [1].

In summary, compared to early applications of AI in learning, modern AI-driven systems appear to rely less on cognitive disequilibrium theories, prioritizing efficiency, personalization, and scaffolding over the introduction of controlled disruptions. While early AI-based tutoring systems leveraged cognitive conflict, conceptual change, and error-driven learning to challenge learners' assumptions, contemporary adaptive learning environments often focus on minimizing uncertainty to maintain engagement and optimize learning paths. Furthermore, despite AI's potential to support complex, systemic learning frameworks (for example by modeling interdependent variables, emergent behaviors, and feedback loops) current AI-driven educational platforms do not seem to embrace such an approach. Instead, they primarily function within linear, individualized learning trajectories, missing an opportunity to cultivate a deeper understanding of dynamic, interconnected knowledge structures. Finally, a vast body of research shows that learning is a social exercise, and interaction and collaboration are at the heart of the learning process (e.g., [64, 6]). However, online collaboration has to be facilitated and moderated [65]. Finally, while AI-based learning tools are often deployed in collaborative educational contexts (such as classrooms where students work side-by-side using platforms like Khan Academy, or MOOCs where learners engage in peer discussion and feedback) most current adaptive systems, to the best of our knowledge, remain designed around the individual learner, with limited modelling of peer dynamics or shared reasoning processes.

While this paper analyses only a limited number of AI-based learning systems, these examples appear to reflect a broader trend toward prioritising efficiency, personalisation, and engagement over conceptual struggle and cognitive conflict.

5 The Winter of Our Discontent? A Reflective Discussion

The observed decline in the integration of cognitive disequilibrium theories in modern AI-driven learning suggests that current AI systems prioritize efficiency, personalization, and scaffolding over productive struggle and conceptual conflict. This tendency aligns with broader historical cycles in AI and Human-Computer Interaction (HCI), where AI's aspirations toward automation and frictionless user experiences have often sidelined research on human-centered interaction and learning [60].

If AI's trajectory continues to prioritize minimizing errors and maximizing smooth, adaptive learning paths, we must ask: does this reflect a fundamental shift in how learning itself is conceptualized within AI-driven environments? And if so, does this shift risk neglecting the very mechanisms (cognitive conflict, disruption, and uncertainty) that have been shown to foster deep learning?

These questions resonate with a broader debate in AI and HCI, as highlighted by Grudin [60] and more recently by Nolwenn Maudet [61]. Grudin's work suggests that AI and HCI have historically experienced alternating cycles of dominance, with HCI flourishing in periods when AI's ambitions faltered. AI, with its ambitious goal of replicating human intelligence, has gone through cycles of boom and bust (commonly referred to as AI winters, periods when optimism faded, funding was cut, and progress stalled). These winters left a vacuum that HCI was often able to fill. During AI downturns, resources were, in fact, redirected toward HCI, leading to major breakthroughs such as graphical user interfaces in the 1980s. However, when AI funding surged again, HCI often found itself in the shadow of AI's grand promises.

In a very provocative way, Nolwenn Maudet [61] states:

> *"The goal of many designers, like Greg Brockman, co-founder of Open AI, when he talks about the future of ChatGPT, is therefore to understand and anticipate your intentions, even if you haven't explicitly expressed them* [62]. *Welcome to the ultimate personalized experience, one that makes your life easier by solving it in advance. So, even if the promoters of AI don't say it in these terms, their vision of the future, what they are aiming for, is the death of interaction design, because wanting to anticipate everything ultimately means automating everything, eliminating all interaction."*

If AI is now at its peak, and if, as Maudet provocatively suggests, AI ultimately seeks to eliminate interaction itself, what does this mean for the future of AI-driven learning?

This leads us to a crucial dilemma: should education technology simply follow AI's current trajectory, optimizing for seamless automation, or should we actively intervene to ensure that AI remains an interactive and challenging learning partner? Should HCI researchers and educational technologists push against the prevailing logic of frictionless AI, advocating instead for intentional disruptions that support deeper learning?

5.1 The Unasked Questions of AI-Driven Learning

This dilemma raises further questions about the fundamental goals of AI-driven learning.

If AI is designed to anticipate and solve problems before they arise, does this redefine what it means to learn (and what learners will become)? Should educational AI systems really prioritize ease and personalization at the expense of challenge and struggle?

Personalization, a central promise of AI in education, ensures that learners receive content tailored to their needs. However, excessive adaptation may reduce the kind of friction that fosters deep learning. Just as overfitting in machine learning results in models that fail to generalize, could over-personalization in AI-driven learning environments create learners who struggle to handle unexpected challenges?

An alternative approach would be to design AI not as a frictionless tutor but as an *agent provocateur* (in the pedagogical sense of a designed disruptor): a presence that introduces productive tension and cognitive friction, delays feedback or even presents contradictory perspectives to encourage deeper engagement. This would shift AI from a passive facilitator back to a *Socratic challenger*, making learning an iterative and interactive process rather than an optimized, seamless experience.

Beyond pedagogical considerations, increasing the integration of AI into education raises critical ecological concerns. Large-scale AI systems, particularly those relying on deep learning and natural language processing, require significant computational power, resulting in high energy consumption and carbon emissions [63]. The environmental cost of training and deploying AI models forces us thus to reconsider when and how AI should be used in learning.

If AI-driven education primarily serves to smooth out challenges and eliminate struggle from the learning process, is this a justifiable use of such a resource-intensive technology? Should we not, instead, reserve AI for learning scenarios where it offers clear and irreplaceable benefits, such as personalized accessibility tools for learners with disabilities or adaptive support for complex problem-solving?

Rather than uncritically expanding AI in education, we must ask: *Where does AI genuinely enhance learning, and where might its benefits outweigh its high ecological cost?* If AI is used merely to automate and simplify learning rather than to deepen engagement and critical thinking, we risk not only compromising the quality of education but also contributing to an unsustainable technological race in academia.

Finally, we raise a question about the role of HCI in shaping the future of AI-powered education. If AI continues to dominate, could HCI research intervene to ensure that learning environments remain interactive, explorative, and rich in meaningful friction?

Ultimately, the challenge is not just whether AI *can* be used in education, but *how* it should be designed and used.

6 Conclusions

This paper has examined the diminishing role of cognitive conflict in AI-driven education, where efficiency and personalization increasingly overshadow productive struggle, and has raised questions about the role of HCI and AI in learning.

As some may have noticed, the title of the previous section is a double play on the concept of AI winters and Steinbeck's *The Winter of Our Discontent* [66], serving as both a reflection and a warning. Just as Steinbeck's protagonist wrestled with integrity and ambition, AI in education faces a choice: will it evolve as a tool that deepens learning through meaningful interaction, or will it sacrifice engagement and challenge in pursuit of seamless automation?

Acknowledgments. This paper was written with the support of various tools and feedback processes. AI-assisted tools were used for grammar refinement and structural clarity, while content critiques were provided by human colleagues who reviewed the manuscript before submission. Their insights and discussions were invaluable in strengthening the arguments and presentation of the work. This article is the result of reflection realised in the context of the EPISTEAM project that has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 101129655.

CRediT author statement.
Ines Di Loreto: Conceptualization, Writing – original draft preparation. Writing – review and editing.

References

1. Luckin, R., Holmes, W., Griffiths, M., & Forcier, L. B. (2016). Intelligence Unleashed: An Argument for AI in Education. Pearson Education.
2. Holmes, W., Bialik, M., & Fadel, C. (2021). Artificial Intelligence in Education: Promises and Implications for Teaching and Learning. Center for Curriculum Redesign.
3. VanLehn, K. (2011). The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. Educational Psychologist, 46(4), 197–221. https://doi.org/10.1080/00461520.2011.611369
4. Wang, X., Xu, X., Zhang, Y., et al. (2024). Exploring the impact of artificial intelligence application in personalized learning environments: Thematic analysis of undergraduates' perceptions in China. Humanities & Social Sciences Communications, 11, 1644. https://doi.org/10.1057/s41599-024-04168-x
5. Jain, N., Kachhara, G., Mathur, V., & Degra, L. (2024). Transforming education: The role of AI in personalized learning and data-driven insights. Journal of Informatics Education and Research, 4(3), Article 1532. https://doi.org/10.52783/jier.v4i3.1532
6. Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education—Where are the educators? International Journal of Educational Technology in Higher Education, 16(1), 39. https://doi.org/10.1186/s41239-019-0171-0
7. Easterday, M. W., Rees Lewis, D. G., & Gerber, E. M. (2018). The logic of design research. Learning: Research and Practice, 4(2), 131–160. https://doi.org/10.1080/23735082.2017.1286367
8. Gibson, D., Broadbent, C., Ma, J., & Smith, S. (2023). Learning theories for artificial intelligence promoting learning processes. British Journal of Educational Technology, 54(4), 1130–1145. https://doi.org/10.1111/bjet.13341
9. Holmes, W., & Tuomi, I. (2022). State of the art and practice in AI in education. European Journal of Education, 57(4), 542–570. https://doi.org/10.1111/ejed.12533
10. Piaget, J. (1936). *La naissance de l'intelligence chez l'enfant.* Neuchâtel, Switzerland: Delachaux et Niestlé.
11. Kibler, J. (2011). Cognitive Disequilibrium. In: Goldstein, S., Naglieri, J.A. (eds) Encyclopedia of Child Behavior and Development. Springer, Boston, MA. https://doi.org/10.1007/978-0-387-79061-9_598
12. Piaget, J. (1975). *L'équilibration des structures cognitives : Problème central du développement.* Paris : Presses Universitaires de France.
13. Festinger, L. (1957). *A theory of cognitive dissonance.* Stanford University Press.
14. Berlyne DE. A theory of human curiosity. British Journal of Psychology. 1954;45(3):180–191. doi: 10.1111/j.2044-8295.1954.tb01243.x.
15. Kidd C, Hayden BY. The Psychology and Neuroscience of Curiosity. Neuron. 2015 Nov 4;88(3):449-60. doi: 10.1016/j.neuron.2015.09.010. PMID: 26539887; PMCID: PMC4635443.
16. Kube T. If the discrepancy between expectations and actual information is too large, expectation change decreases - A replication study. J Behav Ther Exp Psychiatry. 2023 Jun;79:101831. doi: 10.1016/j.jbtep.2022.101831. Epub 2022 Dec 11. PMID: 36521199.
17. Chinn, C. A., & Brewer, W. F. (1993). The role of anomalous data in knowledge acquisition: A theoretical framework and implications for science instruction. Review of educational research, 63(1), 1-49. 10.3102/00346543063001001
18. D'Mello, S., Lehman, B., Pekrun, R., & Graesser, A. (2014). Confusion can be beneficial for learning. Learning and Instruction, 29, 153-170. 10.1016/j.learninstruc.2012.05.003
19. Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. Behavioral and Brain Sciences, 36(3), 181–204. https://doi.org/10.1017/S0140525X12000477

20. Tenenbaum, J. B., Kemp, C., Griffiths, T. L., & Goodman, N. D. (2011). How to grow a mind: Statistics, structure, and abstraction. Science, 331(6022), 1279-1285. https://doi.org/10.1126/science.1192788
21. Tenenbaum, J. B., Griffiths, T. L., & Kemp, C. (2006). Theory-based Bayesian models of inductive learning and reasoning. Trends in Cognitive Sciences, 10(7), 309-318. https://doi.org/10.1016/j.tics.2006.05.009
22. Smith, E. M., & Holmes, N. G. (2017). Reinventing lecture demonstrations to promote student engagement and learning. arXiv preprint arXiv:1712.03174. https://arxiv.org/abs/1712.03174
23. McDermott, L. C., & The Physics Education Group. (1996). Physics by Inquiry: An Introduction to Physics and the Physical Sciences. Wiley.
24. Ríos, L., Pollard, B., Dounas-Frazer, D. R., & Lewandowski, H. J. (2018). Pathways to proposing causes for unexpected experimental results. 10.1119/perc.2018.pr.rios
25. Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. Science, 275(5306), 1593–1599. https://doi.org/10.1126/science.275.5306.1593
26. Grill, F., Guitart-Masip, M., Johansson, J., Stiernman, L., Axelsson, J., Nyberg, L., & Rieckmann, A. (2024). Dopamine release in human associative striatum during reversal learning. Nature Communications, 15(1), Article 59. https://doi.org/10.1038/s41467-023-44358-w
27. Wang, Y., Nassar, M. R., & Schoenbaum, G. (2024). Nucleus accumbens dopamine release reflects Bayesian inference about uncertainty. The Journal of Neuroscience, 44(35), 6662–6675. https://doi.org/10.1523/JNEUROSCI.0120-24.2024
28. Dörner, D. (1996). The logic of failure: Recognizing and avoiding error in complex situations. Basic Books.
29. Sterman, J. D. (2000). Business dynamics: Systems thinking and modeling for a complex world. Irwin/McGraw-Hill.
30. Meadows, D. H. (2008). Thinking in systems: A primer. Chelsea Green Publishing.
31. Capra, F., & Luisi, P. L. (2014). The systems view of life: A unifying vision. Cambridge University Press. 10.1017/CBO9780511895555
32. Feigenbaum, E. A. (1977). The art of artificial intelligence: Themes and case studies of knowledge engineering. Proceedings of the Fifth International Joint Conference on Artificial Intelligence (IJCAI-77), 1014-1029.
33. Buchanan, B. G., & Shortliffe, E. H. (1984). Rule-based expert systems: The MYCIN experiments of the Stanford heuristic programming project. Addison-Wesley.
34. Shortliffe, E. H. (1976). Computer-based medical consultations: MYCIN. Elsevier.
35. Clancey, W. J. (1987). Knowledge-based tutoring: The GUIDON program. MIT Press.
36. Wenger, E. (1987). *Artificial Intelligence and Tutoring Systems: Computational and Cognitive Approaches to the Communication of Knowledge*. Morgan Kaufmann.
37. Anderson, J. R., Boyle, C. F., & Reiser, B. J. (1985). Intelligent tutoring systems. Science, 228(4698), 456-462.
38. Carbonell, J. R. (1970). AI in CAI: An artificial-intelligence approach to computer-assisted instruction. IEEE Transactions on Man-Machine Systems, 11(4), 190-202.
39. Brown, J. S., Burton, R. R., & de Kleer, J. (1982). Pedagogical, natural language, and knowledge engineering techniques in SOPHIE I, II, and III. In D. Sleeman & J. S. Brown (Eds.), Intelligent Tutoring Systems (pp. 227-282). Academic Press.
40. VanLehn, K., Lynch, C., Schulze, K., Shapiro, J. A., Shelby, R., Taylor, L., Treacy, D., Weinstein, A., & Wintersgill, M. (2005). The Andes physics tutoring system: Lessons learned. International Journal of Artificial Intelligence in Education, 15(3), 147–204.
41. Collins, A., Brown, J. S., & Holum, A. (1991). Cognitive apprenticeship: Making thinking visible. American Educator, 15(3), 6-11.
42. Graesser, A. C., Person, N. K., Harter, D., & the Tutoring Research Group. (2001). Teaching tactics and dialog in AutoTutor. International Journal of Artificial Intelligence in Education, 12, 257–279.
43. Brown, J. S., & Burton, R. R. (1978). Diagnostic models for procedural bugs in basic mathematical skills. Cognitive Science, 2(2), 155-192.
44. Johnson, W. L., & Rickel, J. (1997). Steve: An animated pedagogical agent for procedural training in virtual environments. ACM SIGART Bulletin, 8(1-4), 16-21.
45. Jong, T., & van Joolingen, W. R. (1998). Scientific discovery learning with computer simulations. Review of Educational Research, 68(2), 179-201.

46. Tenenbaum, J. B., Kemp, C., Griffiths, T. L., & Goodman, N. D. (2006). How to grow a mind: Statistics, structure, and abstraction. Science, 331(6022), 1279-1285
47. Sailer, M., Hense, J., Mayr, S. K., & Mandl, H. (2017). How gamification motivates: An experimental study of the effects of specific game design elements on psychological need satisfaction. Computers in Human Behavior, 69, 371–380. https://doi.org/10.1016/j.chb.2016.12.033
48. De Jong, T., Sotiriou, S., & Gillet, D. (2014). Innovations in STEM education: The Go-Lab federation of online labs. Smart Learning Environments, 1(1), 3. https://doi.org/10.1186/s40561-014-0003-6
49. Hamari, J., Koivisto, J., & Sarsa, H. (2014). Does gamification work?—A literature review of empirical studies on gamification. Proceedings of the 47th Hawaii International Conference on System Sciences, 3025–3034. https://doi.org/10.1109/HICSS.2014.377
50. Sailer, M., & Homner, L. (2020). The gamification of learning: A meta-analysis. Educational Psychology Review, 32, 77–112. https://doi.org/10.1007/s10648-019-09498-w
51. Seaborn, K., & Fels, D. I. (2015). Gamification in theory and action: A survey. International Journal of Human-Computer Studies, 74, 14–31. https://doi.org/10.1016/j.ijhcs.2014.09.006
52. Issenberg, S. B., McGaghie, W. C., Petrusa, E. R., Gordon, D. L., & Scalese, R. J. (2005). Features and uses of high-fidelity medical simulations that lead to effective learning: A BEME systematic review. Medical Teacher, 27(1), 10–28. https://doi.org/10.1080/01421590500046924
53. Dismukes, R. K., Berman, B. A., & Loukopoulos, L. (2017). The limits of expertise: Rethinking pilot error and the causes of airline accidents. Routledge.
54. Van Merriënboer, J. J. G., & Kirschner, P. A. (2018). Ten steps to complex learning: A systematic approach to four-component instructional design (2nd ed.). Routledge.
55. Roll, I., & Wylie, R. (2016). Evolution and revolution in artificial intelligence in education. International Journal of Artificial Intelligence in Education, 26(2), 582-599. https://doi.org/10.1007/s40593-016-0110-3
56. Luckin, R. (2017). Machine learning and human intelligence: The future of education for the 21st century. UCL Institute of Education Press.
57. Lee, Y., Son, K., Kim, T. S., Kim, J., Chung, J. J. Y., Adar, E., & Kim, J. (2024). One vs. Many: Comprehending Accurate Information from Multiple Erroneous and Inconsistent AI Generations. Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency (FAccT '24), 14 pages. https://doi.org/10.1145/3630106.3662681
58. Molnar, C. (2022). Interpretable machine learning: A guide for making black box models explainable. Leanpub.
59. Shin, D. (2021). The effects of explainability and causability on perception, trust, and acceptance: Implications for explainable AI. International Journal of Human-Computer Studies, 154, 102685. https://doi.org/10.1016/j.ijhcs.2020.102551
60. Grudin, J. (2009). AI and HCI: Two fields divided by a common focus. AI Magazine, 30(4), 48–57. https://doi.org/10.1609/aimag.v30i4.2271
61. Maudet, N. (2024, September 5). Algorithms vs interactions: The revenge of artificial intelligence over design. Institute of Network Cultures. https://networkcultures.org/longform/2024/09/05/algorithms-vs-interactions-the-revenge-of-artificial-intelligence-over-design/
62. The Inside Story of ChatGPT's Astonishing Potential", TED Talk of Greg Brockman in April 2023. Available at : https://www.youtube.com/watch?v=C_78DM8fG6E
63. Strubell, E., Ganesh, A., & McCallum, A. (2019). Energy and policy considerations for deep learning in NLP. Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, 3645–3650.
64. Jonassen, D., Davidson, M., Collins, M., Campbell, J., & Haag, B. B. (1995). Constructivism and computer-mediated communication in distance education. American Journal of Distance Education, 9(2), 7–25. https://doi.org/10.1080/08923649509526885.
65. Salmon, G. (2000). E-moderating - the key to teaching and learning online, (1st ed.,). London: Routledge.
66. Steinbeck, J. (2008). *The winter of our discontent* (S. Shillinglaw, Intro.). Penguin Classics.