As digital technologies assume ever more decision-making power in our homes, schools, and workplaces, learning to program early is no longer an elective skill but a prerequisite for agency. Coding builds the habits of abstraction, logical sequencing, and systematic debugging that underlie organised thinking and creative problem-solving; Seymour Papert famously argued that writing programs lets children “think about thinking itself,” cultivating articulation, persistence, collaboration, and other life-wide competencies [1]. Contemporary studies confirm that even pre-schoolers who tinker with tangible or screen-based code exhibit stronger reasoning and resilience when facing novel tasks [2]. To translate these benefits into the classroom, educators have begun wrapping curricular goals inside playful “escape” scenarios—e.g., *Code-to-Escape* maths quests, AI-literacy puzzles, or cybersecurity breakouts—so that mastering a concept is literally the key that opens the next lock. Building on that movement, we propose an integrated Robot Escape Room for students in school that fuses three research-backed mechanics into one seamless adventure.

A diagram of a puzzle game

AI-generated content may be incorrect.

To translate the three robot-mediated studies into a purely game-based “escape-room” experience, picture a sequence of three locked chambers—each built around a puzzle that mirrors a proven learning mechanic while dispensing with the physical robot: (1) drawing on findings that peer-framed block-coding sped children’s maze completion whereas teacher-style scaffolding yielded deeper knowledge gains [3], the first room’s floor-grid is sealed until teams drag-and-drop Scratch-like blocks to steer a virtual avatar to a goal, with optional “instructor hints” that cost time but boost concept mastery; (2) inspired by the museum exhibit where 109 youngsters improved a chatbot by feeding it better examples and saw its confidence bar rise in real time [4], the second door remains shut until players debug an AI terminal—reinforcing the ideas that humans supply data, algorithms err, and careful curation fixes them; (3) echoing the story-driven cybersecurity curriculum that increased elementary learners’ password-safety scores and computing interest [5], the final vault presents an interactive fairy-tale firewall: learners must spot phishing clues, craft strong pass-phrases, and choose safe narrative branches to collect three cipher digits. Entering the combined code pops the exit, delivering the same blend of rapid engagement and measurable knowledge gain reported by the original research—now entirely through puzzles, no robots required.

**Room 4: AI-Aware Puzzle – Designing for Cognitive Integrity in the Age of LLMs**

1. **Theoretical Challenge:**

With the increasing ubiquity of large language models (LLMs) such as ChatGPT, DeepSeek, and Claude in education, there arises a critical tension between access to intelligent support and preservation of deep cognitive engagement. In game-based learning environments—particularly puzzle or escape-room formats—there is a growing risk that students may circumvent authentic learning processes by relying on LLMs for immediate answers. This undermines opportunities for concept synthesis, higher-order thinking, and problem-solving skill development.

1. **Pedagogical Contribution:**

This puzzle room is deliberately designed to expose and counteract overreliance on LLMs by embedding friction and requiring interpretive reasoning, multi-source verification, and constructive synthesis. While students may consult an embedded AI tool for hints, the information provided is purposefully incomplete, metaphorical, or distributed across fragmented sources. Learners must evaluate the relevance, credibility, and coherence of these fragments, and integrate them into a broader conceptual framework to unlock the final challenge.

A poster of a classroom

AI-generated content may be incorrect.

1. **Instructional Design Features:**
2. LLM agents within the room produce non-directive prompts (e.g., analogies, ambiguous references, partial truths).
3. Clues require cross-checking against in-game artifacts, databases, or prior rooms' content.
4. Correct solutions emerge from student-constructed logic, not single-response answers.
5. A synthesis board or flow diagram tool may be provided to scaffold sense-making and conceptual mapping.
6. **Research Implication:**

This puzzle operationalizes a key future direction in AI-integrated pedagogy: the design of LLM-resilient educational frameworks that preserve the integrity of inquiry-based learning. It contributes to the development of design patterns where AI serves as a catalyst for deeper cognition, rather than a shortcut to outcomes.

A poster of a classroom

AI-generated content may be incorrect.