

**INTERNATIONAL JOURNAL OF
MODERN EDUCATION
(IJMOE)**www.gaexcellence.com/ijmoe**WHERE DOES THE ETHICS LIVE? MECHANIC-CONCEPT
ALIGNMENT AND THE COST OF PARTICIPATION IN
GAME-BASED AI LITERACY OUTREACH**

Sharunizam Shari¹, Wan Ainol Mursyida Ahmad Tarmizi^{2*}, Nor Azlina Azmi³, Mohd Shamsul Daud⁴

¹Faculty of Information Science, Universiti Teknologi MARA Kedah Branch, Malaysia

 nizam057@uitm.edu.my

 <https://orcid.org/0000-0003-0052-8174>

²Faculty of Information Science, Universiti Teknologi MARA Kedah Branch, Malaysia

 wainolmursyida@uitm.edu.my

 <https://orcid.org/0009-0004-9756-6808>

³Faculty of Information Science, Universiti Teknologi MARA Kedah Branch, Malaysia

 nazlina@uitm.edu.my

 <https://orcid.org/0009-0000-5603-7839>

⁴Faculty of Information Science, Universiti Teknologi MARA Kedah Branch, Malaysia

 shamsul559@uitm.edu.my

 <https://orcid.org/0009-0006-4928-4907>

*Corresponding Author

Article Info:**Article history:**

Received date: 30.07.2026

Revised date: 24.08.2026

Accepted date: 17.09.2026

Published date: 28.09.2026

To cite this document:

Shari, S., Tarmizi, W. A. M. A., Azmi, N. A., & Daud, M. S. (2026). Where Does the Ethics Live? Mechanic-Concept Alignment and The Cost of Participation in Game-Based AI Literacy Outreach. *International Journal of Modern Education*, 8(31), 1629-1651.

Abstract:

Game-based activities are increasingly used in public artificial intelligence (AI) literacy outreach. However, limited guidance exists on how game mechanics can meaningfully convey AI ethics concepts rather than simply making an event enjoyable. This study develops a five-level Mechanic-Concept Alignment framework that classifies an activity according to what its rules actually require players to do. Each activity is assessed against three tests: whether play requires handling AI content, whether the rules model the target concept, and whether play shows participants how they themselves performed. The framework was applied to seven activities at a one-day AI ethics outreach event in Kedah, Malaysia. Ninety-eight members of the public attended, of whom 62 completed an evaluation form, producing 124 open-text responses. What participants remembered was shaped more by participation cost, meaning the time, effort, equipment, visibility and adult permission an activity demanded, than by how well its rules carried AI content. Of the 24 respondents who named a specific activity, 18 named one of the least demanding ones. The activity whose rules best modelled an AI concept was named by 8.1 per cent of respondents, compared with 1.6 per cent for activities that only displayed AI content. Comparing the approved plan with the delivered programme explained why. The activity carrying the strongest AI content had been planned

with a loudspeaker but was delivered with headphones, which made it invisible and inaudible to passers-by, limited it to one user at a time, and required parental permission for children. Raising the AI content of an activity therefore tends to raise its cost of entry. To help designers keep ethics content while lowering barriers to taking part, the study offers a one-minute activity model, seven redesign specifications and six practical rules.

DOI: 10.35631/IJMOE.831095

Keywords:

AI Literacy, AI Ethics, Game-Based Learning, Intrinsic Integration, Free-Choice Learning, Outreach Design



© The authors (2026). This is an Open Access article distributed under the terms of the Creative Commons Attribution (CC BY NC) (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact ijmoe@gaexcellence.com.

Introduction

Game-based activities are increasingly used when universities are asked to raise public awareness of artificial intelligence. The reasoning is sound. The subject is abstract, the audience is unselected, attendance is voluntary, and play lowers the barrier to a technical topic in a way that a lecture cannot. Across public libraries in North America, Huang, Gasco-Hernandez, Prasad and Gil-Garcia (2024) catalogued 105 AI-related events and found hands-on workshops and activity formats to be a dominant mode. The same approach is common in Malaysian higher education, where service-learning coursework routinely produces booth-based, activity-driven community events on digital and AI themes.

What designers do not have is guidance at the level where design decisions are actually made, which is the level of the mechanic. A programme team choosing between a matching game, a charades game and a listening game is choosing between rule systems, and rule systems are not neutral carriers of content. A memory-pairs game teaches memory for pairs. Whether it also teaches anything about AI depends entirely on whether the concept has been built into the rules or merely printed on the cards. The learning sciences settled this question two decades ago for classroom games, under the heading of intrinsic integration (Habgood & Ainsworth, 2011), but the principle has not migrated into AI literacy outreach practice, where activity selection is typically driven by material cost, expected popularity and facilitator confidence.

The gap matters because AI ethics is the hardest competency to deliver and the one most often lost. Casal-Otero et al. (2023) found ethics to be the least operationalised of the four AI literacy domains in school provision, most often appended as concluding discussion rather than designed into activity. Zhou, Li, Chai and Chiu (2025) report the same pattern for competency assessment. At field level, Ramírez-Correa, Grandón and Mariano (2026) find that ethical foundations absorb 40 per cent of published generative AI literacy scholarship while evaluation

absorbs only 15 per cent and identify a persistent disconnect between expert ethical frameworks and users' actual ethical awareness. Principles are abundant. Mechanisms for transmitting them to a member of the public standing in a park for ninety seconds are not.

The problem this study addresses operates at three connected levels.

At the level of practice, outreach teams select activities using criteria that have nothing to do with whether the activity can transmit the intended content. Material cost, expected popularity, facilitator confidence and available equipment govern the choice, because no accessible criterion exists for the prior question of whether a given rule system is capable of carrying AI ethics at all. Teams consequently invest equally in activities that transmit the content and activities that cannot and possess no means of telling the two apart before delivery or after it. At the level of theory, the principle that ought to resolve this was established under conditions that do not obtain in outreach. Intrinsic integration tells designers that learning content should be embedded in the rule system rather than wrapped around it, and Habgood and Ainsworth (2011) demonstrated the resulting learning advantage. Their participants, however, and those of the studies that followed, were school pupils in classrooms who were going to play something regardless. Public outreach removes that guarantee. A member of the public in a park may decline any activity at no cost, which means the prior question is not whether an integrated mechanic teaches better but whether anyone will choose it. Existing theory is silent on this, and designers applying it in free-choice settings therefore work without knowing whether its central recommendation transfers.

At the level of measurement, the failure is invisible and consequently self-perpetuating. Outreach programmes are evaluated almost exclusively through post-event satisfaction instruments, which return uniformly high scores and cannot discriminate between an activity that conveyed a concept and one that was merely enjoyable. Ramírez-Correa et al. (2026) locate the same imbalance across the field, with evaluation attracting less than half the scholarly attention devoted to ethical frameworks. A programme team reading only satisfaction means and uptake counts receives no signal that its ethics content failed to land, and will repeat the design.

The cumulative consequence is that public investment in AI ethics literacy cannot presently be steered, because neither designers before delivery nor evaluators after it can distinguish an activity that carries the content from one that decorates it. Addressing this requires a criterion applied at the level of the mechanic, and an account of how that criterion behaves when participation is voluntary.

The aim of this study is to establish which properties of a game mechanic determine whether it can carry AI ethics content in a free-choice public setting, and to derive design guidance from that determination. Three research questions follow.

1. RQ1: How can game mechanics used in AI literacy outreach be classified by the degree to which they carry, rather than decorate, AI ethics content?
2. RQ2: When such mechanics are deployed in a free-choice public setting, what relationship holds between alignment and participant uptake?
3. RQ3: What design changes would raise alignment without incurring the participation costs that appear to suppress it?

The empirical material is a set of seven activities delivered to 98 members of the public at a one-day outreach event in Sungai Petani, Kedah, in January 2026. The study is explicit at the outset about what this material can and cannot support. It is not an effectiveness trial. There was no comparison condition, no pre-test and no per-activity learning measure. It is a design analysis, in which complete documentation of seven contrasting mechanics is examined against a theoretical framework, with field data serving to corroborate rather than to establish. The value lies in the contrast between the seven, not in the outcome of any one.

The study makes five contributions. The principal contribution is theoretical. The study identifies a limit to the established principle of intrinsic integration, which holds that learning content should be built into a game's rules rather than wrapped around them. In free-choice settings, the design moves that build AI content more firmly into the rules also tend to make an activity more demanding to join, and that demand decides participation first, because nobody is obliged to take part. This limit cannot be observed in the classroom settings from which the principle derives, where attendance is compulsory, and it reframes the design objective from achieving the strongest possible alignment to achieving the strongest alignment for each unit of effort asked of the participant.

The second contribution is conceptual. The Mechanic-Concept Alignment framework operationalises intrinsic integration for AI literacy, and in doing so adds a criterion absent from the original formulation: for risk-related content, a mechanic should return evidence of the participant's own fallibility. This evidence test defines the framework's highest level and is warranted by meta-analytic findings that feedback measurably improves synthetic media detection where assertion does not (Diel et al., 2024).

The third contribution is empirical and concerns where participation cost is incurred. By comparing an approved design specification against a delivered programme, the study shows that the decisive costs entered at delivery rather than at design: an equipment substitution unremarkable as logistics disabled the activity best equipped to teach AI risk, and a good-faith content addition raised one activity's alignment and its literacy load simultaneously. Equipment lists and on-the-day content decisions are shown to be pedagogical choices with the same standing as the original specification, and are currently governed as neither.

The fourth contribution is a within-case demonstration, claimed modestly. Outreach is conventionally analysed at programme level through aggregate satisfaction. This study analyses seven contrasting mechanics within a single event, and shows two activities whose stated objectives and transmitted content moved in opposite directions between approval and delivery. Participant recall data corroborate the pattern but cannot establish it, and the finding is offered as an existence proof and an illustrative case rather than as a generalisable effect.

The fifth contribution is practical. The study specifies a one-minute activity model, termed here the one-minute evidential unit, in which a participant makes one judgement about AI content, learns immediately whether it was correct, and adds that result to a public tally. This is accompanied by seven activity-level redesign specifications and six transferable design rules, each costed and implementable within the resources typically available to student-led and library-led AI literacy programming.

The study does not claim novelty for intrinsic integration itself, for the observation that engagement and learning diverge in gamified settings, for the under-design of ethics in AI literacy provision, or for the near-chance accuracy of unaided deepfake detection. Each of these

is established. The novelty lies in their junction with the participation-cost variable, which the outreach setting makes visible.

Literature Review

The review proceeds in four steps. It first establishes that AI ethics is the least operationalised of the AI literacy domains in current provision, then turns to the learning sciences for the principle that explains why, namely intrinsic integration. It next considers what changes when that principle is applied in free-choice public settings rather than in classrooms and closes with the one AI ethics topic whose evidence base is firm enough to serve as a design benchmark. The Mechanic-Concept Alignment framework proposed in the final subsection follows from these four strands.

The Ethics Deficit in AI Literacy Activity Design

Long and Magerko (2020) defined AI literacy as the set of competencies enabling a person to evaluate AI technologies critically, communicate with them effectively, and use them as everyday tools. Ng, Leung, Chu and Qiao (2021) consolidated the field into four domains: knowing and understanding AI, using and applying AI, evaluating and creating with AI, and AI ethics. The UNESCO AI Competency Framework for Students (UNESCO, 2024) places a human-centred mindset and the ethics of AI as the first two of its four dimensions, ahead of techniques and system design, and Malaysia's National Guidelines on AI Governance and Ethics direct comparable expectations at end users (Ministry of Science, Technology and Innovation, 2024).

Provision does not match this ordering. Reviews consistently find ethics under-designed relative to tool use (Casal-Otero et al., 2023; Zhou et al., 2025). Cox (2024), writing from within library and information science, traces a lineage from algorithmic literacy through AI literacy to responsible generative AI literacy, arguing that generative capability makes critical engagement with misinformation and harm a defining rather than peripheral competence. Annapureddy, Fornaroli and Gatica-Pérez (2025) distribute ethical considerations across all twelve of their generative AI literacy competencies rather than isolating them in one, implying that ethics should be a property of every activity rather than the subject of a single session.

The practical question this raises has not been answered: what does it mean, concretely, for a booth activity to have ethics designed into it?

Intrinsic Integration and the Chocolate-Covered Broccoli Problem

The learning sciences supply the analytic vocabulary. Bruckman's (1999) characterisation of educational software as chocolate-covered broccoli names the failure mode, in which the game is used as a sugar coating or reward wrapped around content that remains separate from it. Habgood, Ainsworth and Benford (2005) argued that what matters is not the fantasy context but the rule system, since a fantasy can be exchanged for another without altering what the player actually does. Habgood and Ainsworth (2011) tested the resulting principle of intrinsic integration directly. Intrinsic integration means that the learning content is built into the rules of play itself, so that playing well requires doing the thing being taught, rather than the content being attached to a game that would work equally well without it. Comparing intrinsic, extrinsic and control variants of the same game, they found, both superior learning and greater

free-choice time on task for the intrinsic version. Cutting and Iacovides (2022), in a pre-registered study with 210 participants, identified the mechanism as attentional: players attend to the features the game task requires and ignore the rest.

That mechanism is the crux for outreach design. If players attend only to what the rules require, then an activity whose rules require sorting-coloured beads onto a bracelet has directed attention at beads, whatever the facilitator says while they do it. The broader gamification literature is consistent with this reading. Hamari, Koivisto and Sarsa (2014) found motivational effects common and learning effects conditional on alignment. Zeng, Sun, Looi and Fan (2024), meta-analysing experimental comparisons of gamified and non-gamified instruction, reported a moderate pooled effect (Hedges's $g = 0.782$) with strong moderation by game element and context, together with a novelty effect that decays where designs remain static.

Almost all of this evidence comes from classroom or laboratory settings. That limitation matters more than it appears.

Free-Choice Settings and the Cost of Participation

Falk and Dierking's (2000) contextual model describes learning in museums, parks and festivals as free-choice: learner-directed, socially mediated, non-linear and terminable at will. In such settings the participant pays a cost to engage, denominated in time, attention, social exposure and sometimes equipment, and may decline at any moment without consequence. This introduces a variable that classroom studies of intrinsic integration cannot observe. In a classroom, attendance is compulsory, so a well-integrated but demanding activity is simply a demanding activity and its integration pays off. In a park, a demanding activity is one that people walk past. Intrinsic integration and participation cost are therefore not independent design parameters in free-choice settings, and the relationship between them has, to the authors' knowledge, not been examined.

Synthetic Voice as a Limiting Case

One AI ethics topic has both an unusually firm evidence base and unusually favourable mechanical properties. Diel et al. (2024), pooling 137 effects from 56 studies and 86,155 participants, found overall human deepfake detection accuracy of 55.54 per cent with a confidence interval crossing chance, and audio accuracy of 62.08 per cent with a wide interval. Groh, Sankaranarayanan, Singh, Kim, Lippman and Picard (2024) found limited discernment of fabricated political speech across five pre-registered experiments with 2,215 participants, and Frank et al. (2024) reported comparable results across countries.

Two design properties follow. First, the claim that synthetic media are difficult to detect need not be asserted, because it can be demonstrated on participants themselves in under a minute, most of whom will fail. Second, and decisively, Diel et al. (2024) found feedback training among the few interventions that reliably improve detection, raising accuracy to approximately 65 per cent. A mechanic that tells participants immediately whether they were correct is therefore not merely aligned with the concept but instrumentally effective in a way that exposition cannot be. Synthetic voice discrimination thus represents a natural upper bound for what an outreach mechanic can achieve.

The Mechanic-Concept Alignment Framework

The preceding literature identifies an unsolved problem: ethics is under-designed in AI literacy activity, the learning sciences explain why through intrinsic integration, and free-choice settings introduce a participation constraint that classroom evidence cannot address. This study responds by proposing a classification of AI literacy activities based on three sequential tests applied to the rule system alone, disregarding what facilitators say alongside it.

The content test asks whether completing the activity requires the participant to handle AI-related content. If not, the activity is decorative with respect to AI, whatever its other merits. The structure test asks whether the rule system models the target concept, such that playing correctly requires behaving in the way the concept describes. Printing AI tool names on cards fails this test; requiring a player to specify an instruction precisely enough for a partner to act on it passes it. The evidence test asks whether play generates evidence about the participant's own capability or fallibility and returns it to them as feedback. This is the strongest condition, because it converts an assertion about AI risk into a personal result.

Applying the tests in sequence yields the five levels presented in Table 1.

Table 1: The Mechanic-Concept Alignment (MCA) Framework

Level	Label	Definition	Tests passed	What the rules teach
L0	Decorative	Play involves no AI content; AI appears only in framing or signage	None	Nothing about AI
L1	Adjacent	An AI tool is operated as an instrument, but the object of play is something else	Content test, partially	That an AI tool exists and can be used
L2	Referential	AI content is the material manipulated, but the rules model no AI concept	Content test	Vocabulary, categories, associations
L3	Structural	The rule system enacts the target concept; playing correctly requires performing it	Content and structure tests	The concept, procedurally
L4	Evidential	Play generates and returns evidence of the participant's performance against an AI-related risk	Content, structure and evidence tests	The concept, and their own susceptibility to it

Source: Authors' Analysis

Two clarifications are required. First, the levels are not a quality ranking of activities as experiences. An L0 craft activity may be excellent at drawing families to a booth, which is a legitimate function. The levels rank only the capacity of the mechanic to transmit AI content.

Second, the ladder is orthogonal to the Ng et al. (2021) domains. An activity may target AI ethics at L2, by naming harms, or at L4, by making participants experience their susceptibility to one, and the difference between those two is the subject of this study.

Third, and importantly for interpretation, the framework classifies mechanics rather than booths. A booth comprises a rule system, a facilitator, materials and a social encounter, and only the first of these is classified here. Facilitator explanation in particular operates as a parallel channel that may convey AI content irrespective of what the rules require. That channel is real, and in booth-based outreach it is often substantial, but it is not measurable from the records available in this or most comparable cases, and it is deliberately bracketed so that the property under examination remains constant across activities. An L0 classification therefore indicates that a rule system carries no AI content. It is not a judgement of an activity's value, of the quality of its facilitation, or of what participants took from the encounter as a whole.

Methodology

This section sets out how the seven activities were classified and how that classification was checked against independent evidence. It describes the research design and the case, identifies the four data sources and the order of authority assigned to them, specifies the procedure used to apply the three tests, and explains the coding of participant responses and the construction of the participation-cost audit. It closes by stating plainly what the available data can and cannot establish, and by declaring the position of the research team in relation to the programme studied.

Research Design

This study is a design analysis of a bounded set of contrasting mechanics, conducted in the tradition of design case reporting rather than of controlled evaluation. The unit of analysis is the activity, not the participant and not the programme.

Case and Setting

The setting was a one-day public AI ethics outreach event held on 3 January 2026 at a public park in Sungai Petani, Kedah, Malaysia, delivered by 28 students of a Diploma in Information Management as part of a community outreach course, in partnership with a municipal council, a state public library corporation and a university library. Seven activities ran concurrently from booths in the open park between 09:00 and 16:30. Registration recorded 98 participants aged 10 to 60 years, with a median age of 21.

The case suits the research question for three reasons. The seven activities were designed independently by different sub-teams and therefore vary substantially in mechanic type. Complete design documentation exists for each, including stated objectives, materials and assigned facilitators. An expository component consisting of a lecture and a forum was cancelled on the day, which means the mechanics were the sole channel through which AI content reached participants. The last point makes the setting an unusually clean, if unintended, test of activity-only delivery.

Data Sources

Four sources were used, in a defined order of authority.

The implementation report is the operative document for classification. This 47-page record describes each activity as actually delivered, states the objective the team attributed to it after running it, and contains post-event facilitator reflections recording which activities attracted or failed to attract participants and why. Because the framework classifies mechanics as enacted rather than as intended, the delivered account governs.

The approved programme proposal is a pre-implementation draft, whose action plan tables specify for each activity a provisional objective, strategy, materials list and assigned facilitators. It is used here not as a competing classification but as a baseline against which to detect change.

Refinement between approval and delivery is the normal condition of community programmes rather than a departure from good practice. Partners withdraw, venues impose constraints, equipment is reallocated, and teams improve activities in response to the audience in front of them. Such adjustment is expected and often desirable. The analytical opportunity in this case arises not because change occurred but because the team documented both stages fully enough for the change to be traced, which is uncommon and is a precondition of the analysis that follows.

Facilitator confirmation was sought for two activities whose documented descriptions left the operative rule underspecified. In each case the facilitator responsible for that booth was asked directly what participants were required to do and what the materials carried, and their account was adopted. Both confirmations are incorporated into the classifications reported in Table 2 and are not separately annotated there.

Participant salience data comprised 124 open-text responses from 62 respondents, a return of 63.3 per cent against registrations, to two questions: which part of the programme they found most interesting and why, and which they enjoyed more and why. Classification Procedure

Each activity was classified against the three tests using the design documentation alone, before any participant data were examined, so that uptake could not influence alignment coding. Classification was performed on the rule system as documented, not on stated intent. Where documentation stated an analogical purpose that the rules did not enact, the activity was coded on the rules. Each activity was classified against the implementation record, which governs. It was then compared against the approved proposal so that alignment introduced or refined during delivery could be identified.

Salience Coding and Participation-Cost Audit

Open responses were coded for activity mentions using keyword matching in Malay and English, with manual verification of every match. Because respondents frequently named activities elliptically, matching drew on the design documentation to resolve references. A respondent was counted once per activity regardless of how many times they named it across the two questions.

A participation-cost audit was then constructed for each activity across three dimensions derived from the free-choice learning literature and from constraints visible in the documentation: estimated time to complete, access constraints such as equipment dependency and the need for adult permission, and the cognitive or social load the activity imposes, including language and literacy demands and the degree of public exposure required.

Limitations of Inference and Researcher Position

The salience data are corroborative rather than confirmatory. Only 27 of 62 respondents named any activity, responses were frequently terse, no per-booth attendance counts exist, and mentions confound memorability with exposure, since an activity nobody reached cannot be named. Convergence between three independent signals, namely documented mechanic properties, facilitator-reported uptake and participant mentions, is therefore treated as the evidential basis. No significance tests are reported on mention counts, which would imply a precision the data do not possess.

The research team included members of the academic unit responsible for the course under which the activities were delivered. This provides access to complete design documentation that an external analyst would not obtain and also creates a risk of favourable interpretation. Two safeguards were applied: alignment classification was completed before participant data were examined, and every claim about delivery difficulty is sourced to the organisers' written reflections rather than to team recollection.

Results

Results are reported in three parts. The first classifies the seven mechanics against the three tests and examines the activities whose alignment changed between approval and delivery. The second reports what participants named as most interesting, by activity and by alignment level. The third presents the participation-cost audit, which accounts for the pattern the second part describes.

Classification of the Seven Mechanics

Table 2 reports the classification. The seven activities distribute across all five levels, with a single activity reaching L4.

Table 2: MCA Classification of Seven Activities as Delivered, with Change from Proposal

Activity	Objective recorded in the implementation report	Rule system as delivered	Level	Change from proposal
Make Own Keychain and Bracelet	Support activity for young children; report states it is not directly linked to the AI theme	Assemble beads and fittings to a personal design	L0	None
Design a Bookmark Using Canva	Expose participants to AI technology through Canva in editing and producing creative media	Draw randomised constraints by bursting balloons, then produce a	L1	None

Activity	Objective recorded in the implementation report	Rule system as delivered	Level	Change from proposal
		bookmark in Canva		
Blind Box	Test understanding of each AI tool's function and the ability to identify AI applications in real situations	Draw a card and match AI tools to their correct function	L2	None
Describe the Picture	Demonstrate that humans convey meaning without words, whereas AI requires detailed and specific instruction to produce accurate responses	One participant conveys a target to a partner without naming it	L3	AI rationale articulated
Matching Game	Train short-term memory and observation; encourage social interaction and teamwork	Memory pairs on images unrelated to AI	L0	Algorithmic framing not carried into the rule
Newspaper Puzzle	Expose visitors to the benefits and harms of AI through interactive, creative composition	Cut words from newspapers and compose sentences on the benefits and harms of AI	L2	Raised from L0
Listen to the Song	Distinguish an authentic human voice from a voice converted by AI	Hear an authentic and an AI-converted recording and judge which is which	L4	Loudspeaker replaced by headphones

Source: Authors' Analysis

Four activities require comment. Three changed between approval and delivery and one was verified as documented. Together they are the study's clearest evidence that transmission depends on the rule system rather than on the objective a team writes down.

Describe the Picture strengthened. Its provisional objective at proposal stage named communication, listening and teamwork, with no reference to AI, but its rule already required one participant to convey a target to a partner without naming it. By the implementation report the team had articulated what that rule was doing, recording that the activity demonstrates how AI requires more detailed and specific instruction to produce accurate responses, whereas humans interpret context effectively through non-verbal cues. The mechanic was sound from the outset, and the team recognised its AI rationale through delivering it. This is design insight rather than accident, and it is the pattern the framework predicts: a rule that enacts a concept will surface that concept to the people running it.

The Matching Game followed a different path. At proposal stage it carried the subtitle “Algorithm” and the objective that participants would understand how an AI system matches data to a user’s prompt. The strategy field, where a procedure would be specified, restated that proposition rather than setting out a rule capable of enacting it, and the materials were images unrelated to AI. By delivery the objective was recorded as training short-term memory and observation, with reported outcomes of visual memory and problem solving. The activity is classified L0.

This is a common and understandable pattern. Translating an abstract concept such as algorithmic matching into a rule that a passer-by can execute in two minutes is a genuinely difficult design task, and no proposal template in ordinary use prompts a team to check whether it has been achieved. The activity performed well on the objectives it did carry, drawing the second-highest number of participants mentions in the set and delivering the memory, observation and social interaction outcomes its facilitators recorded.

Read together, these two trajectories show that stated objectives and transmitted content can move independently. Where the rule enacted the concept, the AI rationale emerged through delivery and was documented; where the rule did not, the stated AI objective was not sustained. A review of stated objectives alone would have ranked these two activities in the reverse order, which is the practical case for classifying at the level of the mechanic.

A third verification concerned Blind Box, and it confirmed the documented account. The facilitator confirms that the cards carried AI tool logos and names and that participants matched each tool to its correct function, as the implementation report describes. The activity is classified L2: AI content is the material being manipulated, participants must know what each tool is for in order to succeed, but the matching rule itself models no AI process.

The third case concerns cost rather than alignment. Newspaper Puzzle carried no AI objective at proposal stage, which named reading comprehension, sentence construction and fine motor skill, and specified only the cutting and rearranging of newspaper words into new sentences. The AI content was introduced during delivery, when the team required that, composed sentences address the benefits and harms of AI. This was a genuine pedagogical improvement, and it is the clearest instance in the set of a team raising an activity’s alignment on its own initiative during delivery. It also raised the literacy load from arranging words to composing an argument. The two effects arrived together, which is the difficulty the following section examines rather than a criticism of the decision.

Salience by Alignment Level

Table 3 and Figure 1 report unprompted participant mentions by activity and by level.

Table 3: Participant Mentions by Activity and Alignment Level (n = 62 Respondents, 124 Responses)

Activity	Level	Respondents naming it	Percentage
Design a Bookmark Using Canva	L1	10	16.1
Matching Game	L0	6	9.7
Describe the Picture	L3	5	8.1
Make Own Keychain and Bracelet	L0	4	6.5

Activity	Level	Respondents naming it	Percentage
Blind Box	L2	1	1.6
Listen to the Song	L4	1	1.6
Newspaper Puzzle	L2	0	0.0

Source: Authors' Analysis

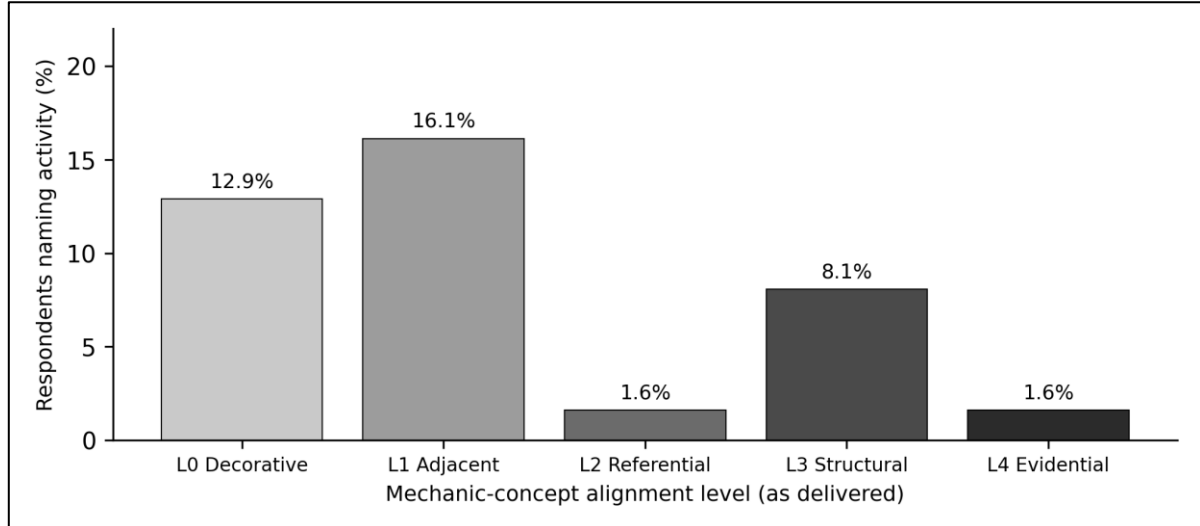


Figure 1: Unprompted Participant Mentions by Mechanic-Concept Alignment Level (n = 62)

Source: Authors' Analysis

Aggregated to level, mentions were 12.9 per cent at L0, 16.1 per cent at L1, 1.6 per cent at L2, 8.1 per cent at L3 and 1.6 per cent at L4. Of the 24 respondents who named any activity, 18 named an activity at L0 or L1 and 7 named an activity at L2 or above.

The pattern is not monotonic in alignment, and the departure is informative. The structurally aligned L3 activity, whose rules enact the concept, outperformed by a factor of five L2 activities, whose rules only present AI content. What distinguishes it from them is not its alignment, which is higher, but its cost, which is lower. Salience in this set therefore tracks participation cost rather than alignment level, and the two variables happen to run together everywhere except at L3.

The two activities designed to deliver ethics content specifically drew one mention between them across 124 responses. Facilitator observation converges independently with this result. The organisers' post-event reflection records that Newspaper Puzzle and Listen to the Song attracted poor participation, attributing this to the activities being perceived as complicated and, for the song activity, to parental reluctance regarding headphone use by children. Neither dataset was collected to answer the present research question, and they agree.

Participation Cost by Activity

Table 4 reports the participation-cost audit. Costs vary widely across the set and are heaviest for the two activities that drew the fewest mentions.

Table 4: Participation Cost by Activity

Activity	Level	Time to complete	Access constraint	Cognitive or social load
Matching Game	L0	2-3 min	None	None
Make Own Keychain and Bracelet	L0	5-10 min	None	None
Design a Bookmark Using Canva	L1	10+ min	Shared devices, limited units	Moderate: design decisions
Blind Box	L2	3-5 min	None	Low: reading required
Newspaper Puzzle	L2	10+ min	None	High: original sentence composition
Describe the Picture	L3	2-3 min	One shared tablet	Moderate: performing in public
Listen to the Song	L4	2-4 min	Headphones, single user	Low

Source: Authors' Analysis

The two activities with zero and near-zero mentions carry the highest aggregate costs, for different reasons. Newspaper Puzzle required a passer-by to cut, arrange and compose an original sentence in a public park, a task with a literacy load closer to homework than to play. Listen to the Song required headphones, which limited throughput to approximately one participant per song, made the activity inaudible and therefore invisible to onlookers, and introduced a permission gate for the children who formed a substantial part of the passing audience.

Notably, this cost was not present in the approved design. The specification for this activity lists a loudspeaker, together with printed lyric excerpts and writing materials, and makes no reference to headphones. The substitution occurred during implementation, and there are several ordinary reasons why it might have: a public park is an acoustically difficult environment, a shared speaker may have been allocated elsewhere, and headphones offer better audio fidelity for a listening task. Each of those is a reasonable operational judgement. The consequence, however, was that an audible, high-throughput, permission-free mechanic became a single-user one that passers-by could neither hear nor observe. The point is not that the decision was careless but that nothing in the planning or approval process would have identified it as a pedagogical decision at all.

Cost does not act alone, however. The most-mentioned activity was not the cheapest. The Canva bookmark demanded around ten minutes and access to a shared device, yet drew 16.1 per cent of mentions, more than any other activity in the set. Its balloon-bursting element supplied noise, spectacle and a reason for bystanders to look. Visibility therefore recruits independently of cost and can offset it, which is the same mechanism operating in reverse at the headphone booth, where a silent activity gave passers-by nothing to notice.

One exception to the cost-alignment correlation appears in the data. Describe the Picture achieved L3 while remaining inexpensive: two to three minutes, one shared tablet, no reading requirement, and a mechanic legible to a five-year-old. It drew 8.1 per cent of mentions, more than either L2 activity, and the organiser report notes that although aimed at children it also attracted teenagers and adults. It's one substantial cost, public performance, is partly self-

financing, because a person miming in a park constitutes the recruitment display that the headphone activity lacked.

Discussion

The discussion interprets these results in four moves. It first accounts for the relationship between alignment, participation cost and participant recall, and identifies the limit this places on the principle of intrinsic integration. It then locates the point in the programme cycle at which participation cost is incurred, sets out redesign specifications for each of the seven mechanics together with a general design target, and draws out the implications for programme designers, for the information professions and for evaluation practice. It closes with the limitations of the study.

The Alignment-Accessibility Tension

The results describe a relationship between participant recall and participation cost that is stronger and cleaner than any relationship between recall and alignment. Alignment and cost happened to run together across six of the seven activities, which would have made the two indistinguishable had the set been uniform. The seventh activity separates them. Describe the Picture occupies the second-highest alignment level at close to the lowest cost, and it outperformed by a factor of five both activities whose rules only present AI content. Cost, not alignment, is doing the work.

Mentions confound memorability with exposure, and the cost audit locates the cause in differential access rather than differential memorability once accessed. The activities that carried AI ethics were not forgotten by those who played them; they were reached by very few people.

This study names the pattern the alignment-accessibility tension, meaning the trade-off between how well an activity's rules carry the concept and how easy that activity is to join. In free-choice settings, the design moves that raise alignment tend also to raise participation cost, and participation cost decides the outcome because attendance is voluntary. The tendency is strong but not inevitable, which is the practically important qualification. Describe the Picture demonstrates that a mechanic can enact a concept in two minutes, with one shared device and no reading requirement, and the design task is to find such mechanics rather than to accept the trade-off as given.

This tension marks a limit on the principle of intrinsic integration, that is, a condition under which the principle does not hold in the way classroom research suggests. Habgood and Ainsworth (2011) established that integrating content into the rule system improves learning and free-choice time on task and Cutting and Iacovides (2022) identified attention as the mechanism. Both findings concern participants who are already playing. Habgood and Ainsworth's free choice operated between two versions of a game children were already engaged with, not between playing and walking away. In outreach settings the prior question is whether anyone plays at all, and the design moves that raise integration frequently raise the threshold for entry by adding reading, equipment, time, performance or permission.

The implication is not that intrinsic integration fails outside the classroom. It is that in free-choice settings integration must be evaluated jointly with participation cost, and that designs maximising integration without regard to cost will be dominated by cheaper, less aligned activities, because the latter reach people. The design objective is therefore not maximal alignment but maximal alignment per unit of participation cost.

This also reframes what appears as failure in outreach reporting. The two activities that failed in this case were the two that took AI ethics most seriously. Their failure was a cost failure rather than a concept failure, and it is recoverable by redesign rather than by abandonment. A programme team reading only uptake figures would draw the opposite conclusion and discard them.

Where Participation Cost Enters

Comparing the approved proposal against the implementation record locates the point at which participation cost is incurred, and in this case, it was not the design stage. Two of the three cost increases arose during delivery, one alongside a genuine pedagogical improvement. Both were the kind of adjustment competent teams make routinely, which is precisely what makes the finding generalisable rather than particular to this programme.

The headphone substitution is the clearest instance. An approved specification listing a loudspeaker was implemented with headphones, which is an unremarkable operational decision, perhaps taken because a park is noisy or because the available speaker was assigned elsewhere. Its consequences were severe and entirely predictable in hindsight: throughput collapsed to one participant per song, the activity became inaudible and therefore invisible to passers-by, and a permission gate appeared for unaccompanied children. The activity best equipped to teach AI risk was disabled by an equipment swap that no approval process would have flagged, because equipment lists are treated as logistics rather than as pedagogy.

The Newspaper Puzzle shows the same mechanism operating in the opposite direction on alignment. Facilitators raised the activity from L0 to L2 during delivery by requiring that composed sentences address the benefits and harms of AI, which was a genuine pedagogical improvement. It simultaneously raised the literacy load from arranging words to composing an argument, and the activity drew no mentions at all. Improving alignment without auditing the cost of that improvement produced an activity nobody completed.

The implication for practice is that participation cost warrants attention at the point of delivery and not only at the point of design. Equipment substitutions and content additions made on the day carry pedagogical consequences comparable to those of the original specification, yet they are typically treated as logistics and are made without a framework that would surface their effects. Supplying such a framework, rather than attributing fault, is the purpose of the classification proposed here.

Redesign Specifications

The practical contribution of this analysis is a set of specifications that move each mechanic up the ladder while holding or lowering its cost. Table 5 states them.

Table 5: Redesign Specifications for the Seven Mechanics

Activity	Current	Target	Specification
Keychain and Bracelet	L0	L1	Print the finished object with a three-question checklist for verifying an AI-generated message. The mechanic is unchanged; the artefact carries content home.
Bookmark (Canva)	L1	L3	Replace free design with a paired task: produce the same bookmark from a one-line brief, then from a ten-word brief, and compare. The rules then enact prompt specificity.
Matching Game	L0	L3	Supply the rule the approved design omitted. Change the match from image-to-image to prompt-to-output, requiring participants to pair an instruction with the result it would produce, including mismatched pairs that model a poorly specified prompt. The stated algorithmic objective is then enacted rather than asserted. Cost unchanged.
Blind Box	L2	L3	Extend tool-to-function matching into task-to-tool-to-limitation triples, requiring participants to select a tool for a stated task and then say what that tool would get wrong. The limitation card carries the ethics. Adds approximately one minute.
Newspaper Puzzle	L2	L3 at lower cost	Abandon free composition. Supply pre-cut phrase strips and require sorting into benefit, harm and depends-on-context columns on a shared public board. Literacy load falls from composition to sorting, time falls below two minutes, and the accumulating board recruits further participants.
Describe the Picture	L3	L4	Retain the mechanic and add a tally: record whether each pair succeeded and post a running success rate beside the booth. Cost: one tally sheet.
Listen to the Song	L4 partial	L4 full at lower cost	Restore the loudspeaker specified in the approved design, playing both versions aloud on a fixed loop; collect votes with two buckets and tokens; display the running error rate on a whiteboard. Throughput rises from one to many, the permission gate disappears, and the collective error rate becomes the argument. This is a reversion to the approved specification with the addition of a public tally, not a new design.

Source: Authors' Analysis

The final specification is the one the external evidence most strongly supports. Diel et al. (2024) establish that unaided detection sits near chance, so a public tally will display a failure rate that no facilitator needs to assert and further establish that feedback training measurably improves detection. A booth that reveals the answer after each round therefore delivers the one intervention with meta-analytic support, at the cost of a whiteboard.

The One-Minute Evidential Unit

Generalising from Table 5, this study proposes a design target for free-choice AI literacy outreach: the one-minute evidential unit. An activity should, within approximately sixty seconds and without equipment that gates access, require the participant to make a judgement about AI-related content, reveal whether they were correct, and contribute their result to a visible public aggregate.

Each element performs specific work. The time bound respects the dwell time free-choice settings afford (Falk & Dierking, 2000). The judgement requirement satisfies the structure test. The reveal satisfies the evidence test and imports the feedback effect documented by Diel et al. (2024). The public aggregate addresses the recruitment problem that suppressed the highest-aligned activity in this case, converting the mechanic from a private transaction into a spectacle, and simultaneously producing the programme's own outcome data at no additional cost.

Design Heuristics

Six heuristics follow for practitioners. First, classify before building: apply the three tests to each proposed activity at proposal stage, and treat any activity failing the content test as a recruitment device rather than as AI literacy provision. Second, audit cost alongside alignment, since time, equipment, literacy, exposure and permission determine who reaches the content. Third, never gate the payload behind scarce equipment, which was the single most consequential error observed in this case. Fourth, make the mechanic audible or visible, because activities producing ambient spectacle recruit and those conducted in private silence do not. Fifth, return the result, since an activity that does not tell participants how they performed forfeits both the learning effect and the evaluation data. Sixth, aggregate in public, converting individual results into a community-level claim about AI risk and yielding an outcome measure without a questionnaire.

Implications for Practice

For programme designers, the framework offers a proposal-stage instrument that costs minutes to apply and surfaces the composition of a programme before it is built. A booth event consisting of six L0 and L1 activities is a public relations exercise with AI branding, which may be defensible but should not be reported as AI literacy delivery.

For the information professions, the analysis bears directly on the brokerage role that public and academic libraries are increasingly assuming for AI awareness (Cox, 2024; Huang, Gascó-Hernandez, & Gil-Garcia, 2025; Huang et al., 2024; Ko & Chiu, 2024; Ru & Tang, 2025). Libraries deliver into precisely the free-choice conditions modelled here: walk-in audiences, voluntary attendance, mixed ages, short dwell times and constrained equipment. The alignment-accessibility tension is therefore a standing design condition of library AI programming rather than an artefact of one event.

For evaluation, an L4 mechanic generates outcome data as a by-product of play, including detection accuracy and error rates by age group. Lintner (2024) found that of 16 validated AI literacy scales only three were performance-based, and subsequent instruments have deliberately shifted toward factual and task-based assessment (Hornberger, Bewersdorff, Schiff, & Nerdel, 2025; Jin, Martínez-Maldonado, Gašević, & Yan, 2025; Soto-Sanfiel,

Angulo-Brunet, & Lutz, 2025). Outreach programmes cannot administer such instruments in a park, but a well-designed evidential booth is a performance measure in disguise.

Limitations of the Study

The salience data cannot establish causation, and the inverse relationship reported here is a description whose explanation rests on the cost audit and facilitator observation rather than on the mentions themselves. No per-booth attendance was recorded, so exposure and memorability cannot be separated. Activity mentions were coded by a single analyst; the coding rules and full response corpus are available for audit, but formal intercoder reliability was not established. The participation-cost audit rests on documented materials and facilitator reports rather than on timed observation, so its time estimates are approximate. Divergence between proposal and delivery was reconstructed from two document sets and one facilitator confirmation rather than observed directly, so the reasons for the equipment substitution and the content addition are inferred from their traces rather than established from the decisions themselves.

Facilitator confirmation was sought for two activities whose documented descriptions left the operative rule underspecified and was not extended to systematic verification of all seven, which replication should include. The framework's rule-only commitment also brackets facilitator explanation, which in booth settings may transmit substantial AI content that no classification here would register; measuring that parallel channel requires observational methods this study did not employ. The framework itself is proposed rather than validated, being derived deductively from intrinsic integration theory and applied to seven cases; its discriminant utility should be tested against a larger and more varied activity corpus. Finally, the redesign specifications in Table 5 are untested hypotheses.

Conclusion

Game-based activities are widely used in public AI literacy work without any accompanying account of which mechanics can carry the content. This study supplies one. The Mechanic-Concept Alignment framework classifies an activity by whether play requires AI content, whether the rules model the concept, and whether play returns evidence of the participant's own performance, yielding five levels from decorative to evidential.

Applied to seven activities delivered to 98 members of the public, the framework showed the set spread across all five levels with only one reaching the highest and revealed an inverse relationship between alignment and participant uptake. The two activities carrying explicit AI ethics content drew one mention across 124 open responses and were independently reported as the least attended. A participation-cost audit locates the cause not in the concepts but in the entry conditions, with the highest-value mechanic gated behind a single pair of headphones. This alignment-accessibility tension is proposed as a limit on the principle of intrinsic integration that settings with compulsory attendance cannot reveal, and constitutes the study's principal novelty: not that content should be built into game rules, which is established, but that in voluntary settings the cost of entering the rules determines whether the content reaches anyone, and that a mechanic returning evidence of the participant's own fallibility is the form of integration best suited to AI risk.

The practical response is not to retreat to decorative activities that recruit well and teach nothing, but to design for maximal alignment per unit of participation cost: one minute, one judgement, one reveal, one public tally. On that specification, the activity that reached almost nobody in this case becomes the centre of the event, its rules unchanged, its headphones replaced by a speaker and a whiteboard.

Further research should proceed through a design-based research cycle in which the redesigned mechanics are delivered and their embedded accuracy data compared across conditions, which the one-minute evidential unit makes feasible at no additional evaluation cost. The framework should also be applied to activity corpora from other institutions and settings to test its discriminant utility, and to school and library programming where dwell time and participation constraints differ from those of an open public park.

-
- Acknowledgements:** The authors would like to express their sincere gratitude to the Research Management Unit, Universiti Teknologi MARA Cawangan Kedah, for the support extended throughout the course of this research. Appreciation is also recorded to the students who designed and delivered the activities analysed here and documented them with a candour that made this analysis possible, and to the municipal council, state public library corporation and university library that partnered in the programme.
- Funding Statement:** No funding. This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.
- Conflict of Interest Statement:** The authors declare that there is no conflict of interest regarding the publication of this paper.
- Ethics Statement:** This study is a secondary analysis of anonymised administrative records generated as routine programme documentation. It did not involve any intervention, experimentation, or direct contact with human participants for research purposes, and therefore did not require institutional ethics review. Consent for documentation was obtained from participants at the point of registration, and participation in the programme was voluntary. All direct identifiers were removed prior to analysis, no individual is identifiable in any reported result, and the data were used solely for academic purposes. The authors confirm that the research was conducted in accordance with accepted academic integrity and ethical publishing standards.
- Author Contribution Statement:** All authors contributed significantly to the development of this manuscript. Author 1 served as the main author and was responsible for the conceptualisation of the study, the development of the analytical framework and the overall supervision of the work. Author 2 was responsible for data preparation, coding and analysis. Authors 3 and 4 contributed to the classification of mechanic alignment levels, the compilation of programme documentation, the drafting of the manuscript and its critical revision. Author 5 served as the corresponding author and was responsible for the overall coordination of the research team, liaison with the programme organisers and partner institutions, and verification of activity details with the facilitators concerned. All authors read the finalised article thoroughly, provided feedback on its content, structure and interpretation, and approved the final version of the manuscript prior to submission.
-

References

- Annapureddy, R., Fornaroli, A., & Gatica-Pérez, D. (2025). Generative AI literacy: Twelve defining competencies. *Digital Government: Research and Practice*, 6(1), 1-21. <https://doi.org/10.1145/3685680>
- Bruckman, A. (1999). *Can educational be fun?* Paper presented at the Game Developers Conference, San Jose, CA.
- Casal-Otero, L., Catala, A., Fernández-Morante, C., Taboada, M., Cebreiro, B., & Barro, S. (2023). AI literacy in K-12: A systematic literature review. *International Journal of STEM Education*, 10, 29. <https://doi.org/10.1186/s40594-023-00418-7>
- Cox, A. (2024). Algorithmic literacy, AI literacy and responsible generative AI literacy. *Journal of Web Librarianship*, 18(2), 93-110. <https://doi.org/10.1080/19322909.2024.2395341>
- Cutting, J., & Iacovides, I. (2022). Learning by doing: Intrinsic integration directs attention to increase learning in games. *Proceedings of the ACM on Human-Computer Interaction*, 6(CHI PLAY), 1-19. <https://doi.org/10.1145/3549503>
- Diel, A., Lalg, T., Schröter, I. C., MacDorman, K. F., Teufel, M., & Bäuerle, A. (2024). Human performance in detecting deepfakes: A systematic review and meta-analysis of 56 papers. *Computers in Human Behavior Reports*, 16, 100538. <https://doi.org/10.1016/j.chbr.2024.100538>
- Falk, J. H., & Dierking, L. D. (2000). *Learning from museums: Visitor experiences and the making of meaning*. AltaMira Press.
- Frank, J., Herbert, F., Ricker, J., Schönherr, L., Eisenhofer, T., Fischer, A., ... Holz, T. (2024). A representative study on human detection of artificially generated media across countries. In *Proceedings of the 2024 IEEE Symposium on Security and Privacy* (pp. 55-73). IEEE. <https://doi.org/10.1109/SP54263.2024.00159>
- Groh, M., Sankaranarayanan, A., Singh, N., Kim, D. Y., Lippman, A., & Picard, R. (2024). Human detection of political speech deepfakes across transcripts, audio, and video. *Nature Communications*, 15, 7629. <https://doi.org/10.1038/s41467-024-51998-z>
- Habgood, M. P. J., & Ainsworth, S. E. (2011). Motivating children to learn effectively: Exploring the value of intrinsic integration in educational games. *Journal of the Learning Sciences*, 20(2), 169-206. <https://doi.org/10.1080/10508406.2010.508029>
- Habgood, M. P. J., Ainsworth, S. E., & Benford, S. (2005). Endogenous fantasy and learning in digital games. *Simulation & Gaming*, 36(4), 483-498. <https://doi.org/10.1177/1046878105282276>
- Hamari, J., Koivisto, J., & Sarsa, H. (2014). Does gamification work? A literature review of empirical studies on gamification. In *Proceedings of the 47th Hawaii International Conference on System Sciences* (pp. 3025-3034). IEEE. <https://doi.org/10.1109/HICSS.2014.377>
- Hornberger, M., Bewersdorff, A., Schiff, D. S., & Nerdel, C. (2025). Development and validation of a short AI literacy test (AILIT-S) for university students. *Computers in Human Behavior: Artificial Humans*, 5, 100176. <https://doi.org/10.1016/j.chbah.2025.100176>
- Huang, Z.-X., Gascó-Hernandez, M., & Gil-Garcia, J. R. (2025). How could public libraries empower communities to participate in government AI initiatives? A practical view of benefits and challenges. In *Proceedings of the 26th Annual International Conference on Digital Government Research*. TU Delft Open. <https://doi.org/10.59490/dgo.2025.1034>
- Huang, Z.-X., Gascó-Hernandez, M., Prasad, A., & Gil-Garcia, J. R. (2024). Public libraries and their role in raising awareness about AI and fostering inclusive civic engagement:

- Current practices and future development. In *Proceedings of the 25th Annual International Conference on Digital Government Research* (pp. 703-710). ACM. <https://doi.org/10.1145/3657054.3657137>
- Jin, Y., Martinez-Maldonado, R., Gašević, D., & Yan, L. (2025). GLAT: The Generative AI Literacy Assessment Test. *Computers and Education: Artificial Intelligence*, 9, 100436. <https://doi.org/10.1016/j.caeai.2025.100436>
- Ko, C. R., & Chiu, M.-H. (2024). How can academic librarians support generative AI literacy: An analysis of library guides using the ACRL information literacy framework. *Proceedings of the Association for Information Science and Technology*, 61, 977-979. <https://doi.org/10.1002/pra2.1159>
- Lintner, T. (2024). A systematic review of AI literacy scales. *npj Science of Learning*, 9(1), 50. <https://doi.org/10.1038/s41539-024-00264-4>
- Long, D., & Magerko, B. (2020). What is AI literacy? Competencies and design considerations. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1-16). ACM. <https://doi.org/10.1145/3313831.3376727>
- Malone, T. W., & Lepper, M. R. (1987). Making learning fun: A taxonomy of intrinsic motivations for learning. In R. E. Snow & M. J. Farr (Eds.), *Aptitude, learning and instruction* (Vol. 3, pp. 223-253). Erlbaum.
- Ministry of Science, Technology and Innovation. (2024). *National guidelines on AI governance and ethics*. Ministry of Science, Technology and Innovation.
- Ng, D. T. K., Leung, J. K. L., Chu, S. K. W., & Qiao, M. S. (2021). Conceptualizing AI literacy: An exploratory review. *Computers and Education: Artificial Intelligence*, 2, 100041. <https://doi.org/10.1016/j.caeai.2021.100041>
- Ramírez-Correa, P., Grandón, E. E., & Mariano, A. M. (2026). Mapping the landscape of generative artificial intelligence literacy: A systematic review toward social, ethical, and sustainable AI adoption. *Sustainability*, 18(3), 1429. <https://doi.org/10.3390/su18031429>
- Ru, K. C., & Tang, R. (2025). Promoting AI literacy through US academic libraries: An analysis of LibGuides from ARL and Oberlin Group libraries using the EDUCAUSE AI literacy framework. *Information Research*, 30, 847-865. <https://doi.org/10.47989/ir30iConf47182>
- Soto-Sanfiel, M. T., Angulo-Brunet, A., & Lutz, C. (2025). The Scale of Artificial Intelligence Literacy for All (SAIL4ALL): Assessing knowledge of artificial intelligence in all adult populations. *Humanities and Social Sciences Communications*, 12. <https://doi.org/10.1057/s41599-025-05978-3>
- UNESCO. (2024). *AI competency framework for students*. UNESCO.
- Zeng, J., Sun, D., Looi, C.-K., & Fan, A. C. W. (2024). Exploring the impact of gamification on students' academic performance: A comprehensive meta-analysis of studies from the year 2008 to 2023. *British Journal of Educational Technology*, 55(6), 2478-2502. <https://doi.org/10.1111/bjet.13471>
- Zhou, X., Li, Y., Chai, C. S., & Chiu, T. K. F. (2025). Defining, enhancing, and assessing artificial intelligence literacy and competency in K-12 education from a systematic review. *Interactive Learning Environments*. Advance online publication. <https://doi.org/10.1080/10494820.2025.2487538>