



PDF Download
3731459.3773305.pdf
15 March 2026
Total Citations: 0
Total Downloads: 140

Latest updates: <https://dl.acm.org/doi/10.1145/3731459.3773305>

RESEARCH-ARTICLE

Designing and Evaluating Museum Exhibit Prototypes to Foster Middle Schoolers' AI Literacy through Creativity and Embodiment

HASTI DARABIPOURSHIRAZ, Northwestern University, Evanston, IL, United States

SOPHIE ROLLINS, School of Communication, Evanston, IL, United States

MILKA TRAJKOVA, Georgia Institute of Technology, Atlanta, GA, United States

CHENGZHI ZHANG, Georgia Institute of Technology, Atlanta, GA, United States

YASMINE BELGHITH, Georgia Institute of Technology, Atlanta, GA, United States

TOM MCKLIN, The Findings Group, LLC, Decatur, GA, United States

[View all](#)

Open Access Support provided by:

[Northwestern University](#)

[Georgia Institute of Technology](#)

[School of Communication](#)

[The Findings Group, LLC](#)

Published: 07 March 2026

[Citation in BibTeX format](#)

TEI '26: Twentieth International
Conference on Tangible, Embedded, and
Embodied Interaction
March 8 - 11, 2026
IL, Chicago, USA

Conference Sponsors:
SIGCHI

Designing and Evaluating Museum Exhibit Prototypes to Foster Middle Schoolers' AI Literacy through Creativity and Embodiment

Hasti Darabipourshiraz*
Computer Science
Northwestern University
Evanston, Illinois, USA
hastidarabipourshiraz2028@u.northwestern.edu

Sophie Rollins*
School of Communication
Northwestern University
Evanston, Illinois, USA
sophie.rollins2023@u.northwestern.edu

Milka Trajkova
Georgia Institute of Technology
Atlanta, Georgia, USA
mtrajkova3@gatech.edu

Chengzhi Zhang
Expressive Machinery Lab
Georgia Institute of Technology
Atlanta, Georgia, USA
czhang694@gatech.edu

Yasmine Belghith
School of Interactive Computing
Georgia Institute of Technology
Atlanta, Georgia, USA
byasmine@gatech.edu

Tom McKlin
The Findings Group
Decatur, Georgia, USA
tom@thefindingsgroup.org

Jessica Roberts
Interactive Computing
Georgia Institute of Technology
Atlanta, Georgia, USA
jessica.roberts@cc.gatech.edu

Brian Magerko
Georgia Tech
Atlanta, Georgia, USA
magerko@gatech.edu

Duri Long
Department of Communication Studies
Northwestern University
Evanston, Illinois, USA
Duri@northwestern.edu

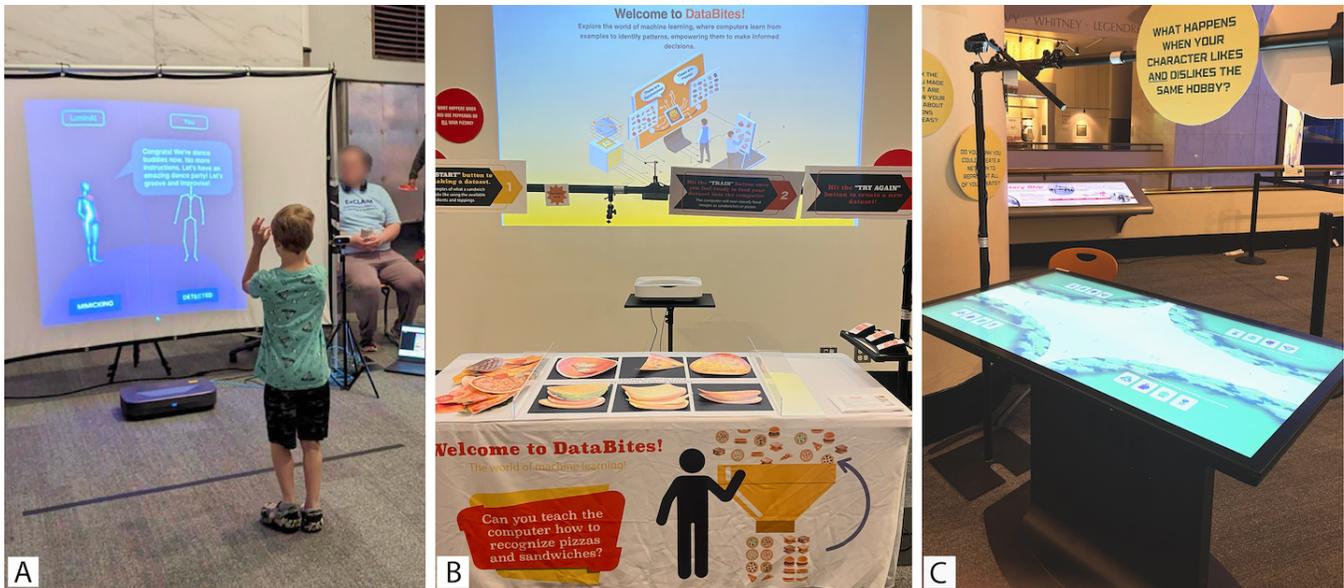


Figure 1: A) *LuminAlx* exhibit B) *DataBites* exhibit C) *Knowledge Net* exhibit

*Both authors contributed equally to this research.



This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.

TEI '26, Chicago, IL, USA

© 2026 Copyright held by the owner/author(s).

ACM ISBN 979-8-4007-1868-7/26/03

<https://doi.org/10.1145/3731459.3773305>

Abstract

Museums play a critical role in promoting public understanding of emerging technologies like artificial intelligence (AI), but it is unclear what design features lead to learning about AI in museums. We contribute a design research exploration of how embodiment and creativity foster AI literacy in museum exhibits. We present design prototypes of three museum exhibits—*DataBites*, *Knowledge Net*, and *LuminAlx*—that aim to teach middle schoolers about AI. We present results from a qualitative analysis of an in-museum study in which we examined participants' understanding of and interest

in AI through interviews and video recordings. Our findings illuminate how creativity fosters interest in AI and how different forms of embodiment contribute to learning about AI. We recommend that AI museum exhibits utilize creative and personally relevant activities to engage middle schoolers, support hybrid conceptualizations of AI, and leverage tangible interaction to make AI concepts approachable.

CCS Concepts

• **Human-centered computing** → **Interface design prototyping**; **Interactive systems and tools**; **Empirical studies in interaction design**; • **Applied computing** → **Interactive learning environments**.

Keywords

informal learning, museum exhibit, AI literacy, creativity, embodiment, prototyping, design research, AI education

ACM Reference Format:

Hasti Darabipourshiraz, Sophie Rollins, Milka Trajkova, Chengzhi Zhang, Yasmine Belghith, Tom McKlin, Jessica Roberts, Brian Magerko, and Duri Long. 2026. Designing and Evaluating Museum Exhibit Prototypes to Foster Middle Schoolers' AI Literacy through Creativity and Embodiment. In *Twentieth International Conference on Tangible, Embedded, and Embodied Interaction (TEI '26), March 08–11, 2026, Chicago, IL, USA*. ACM, New York, NY, USA, 20 pages. <https://doi.org/10.1145/3731459.3773305>

1 Introduction

Museums play a critical role in promoting public understanding of emerging technologies, including artificial intelligence (AI), as they are spaces that encourage inter-generational dialogue, collaborative sense making, and creative, interest-driven exploration [24]. Designing interactive systems to foster broader public *AI literacy*—or the set of competencies necessary for informed, critical engagement with AI [41]—has been a growing area of interest in HCI research [6, 64, 70], especially for young people whose interests and career aspirations are shaped by this evolving technology. However, only a few projects have explored how to design interactive museum exhibits to foster AI literacy [34, 37]. There is a need for more research investigating what design elements support learning about AI through museum exhibits.

Prior work indicates that interactions that involve *embodiment* (moving with the body or using body knowledge [39, 47, 50]) and *creativity* (making novel or personally relevant artifacts [14]) are effective at promoting learning about AI in informal learning spaces [21, 40]; however, this literature has not yet pinpointed specific types or aspects of embodiment and creativity that lead to learning about AI in museum exhibits. We aim to surface such aspects so that exhibit designers can effectively foster learning about and interest in AI through their designs. In this paper, we present the first stage of an iterative design-based research process [1] investigating how embodiment and creativity contribute to learning about AI in museum exhibits. We explore two research questions: RQ1) *How can we leverage embodiment and creativity to design museum exhibits to foster middle schoolers' (ages 10-14) AI literacy?*; and RQ2) *How do embodiment and creativity influence the development of middle schoolers' AI literacy and interest in AI?* To answer these questions,

we engaged in conjecture mapping [58] to articulate hypotheses about how we expect embodiment and creativity to facilitate learning in visitor interactions with museum exhibits, instantiated these conjectures in design prototypes, and evaluated the designs in-situ through a study with middle schoolers in a science and technology museum.

This article presents the designs of three novel museum exhibits that teach middle schoolers about AI: *DataBites* [18], in which learners train a machine learning model to recognize pizzas and sandwiches; *Knowledge Net* [57], in which learners control a virtual world by creating a knowledge representation; and *LuminAIx* [67], in which learners dance with an AI dance partner. We conducted a museum study to determine whether—and how—our designs foster learners' AI literacy and interest in AI. We synthesize our findings into future directions for our designs and design principles for museum exhibits to foster AI literacy. The main contributions of this paper include 1) a set of novel exhibit designs that leverage creativity and embodiment to foster AI literacy, 2) empirical findings and design principles that articulate mechanisms of creativity and embodiment that lead to AI learning and interest development, and 3) a multi-case analysis of different types of embodiment and creativity leveraged in AI learning across three exhibits.

2 Related Work

2.1 AI Museum Exhibits

A few projects have investigated how to foster AI literacy through museum exhibits. Some museums have developed exhibits to engage young learners in conversations about AI by showing applications of AI [2], teaching about data and algorithms [26], and comparing human vs. AI capabilities [4]. However, there have only been a few empirical studies that examined how visitors interact with and learn from AI exhibits [34, 37]. These studies posit that collaboration [34] and personally relevant activities [37] can engage visitors, but more research is needed to investigate the role of embodiment and creativity in learning about AI through a museum exhibit.

More research has investigated how to foster middle schoolers' AI literacy in other informal learning environments, such as summer camps or after-school groups [53, 68, 70], video games [6, 64], and at-home learning [21, 43–45]. However, design features that support learning in these spaces do not necessarily transfer to museums. In museums, interactions are typically brief, collaborative, visitor-led, and need to be enjoyable and rapidly engaging [31], whereas in other environments, learners may be led by an instructor and engage with materials for extended periods of time. More research is needed to investigate whether design features for fostering AI literacy in informal learning spaces (some of which draw on creativity [17, 20, 43, 64] and embodiment [22, 27, 35, 43, 45, 46, 61, 69]) transfer to museums, or whether new approaches are needed in museums.

2.2 Embodiment & Learning

The role that the body, movement, and bodily knowledge play in learning has been explored extensively [39, 47, 50]. Learning scientists have conceptualized various forms of embodiment that contribute to learning in different ways [71]. In our work, we have

selected three forms of embodiment—*tangible user interfaces*, *body knowledge*, and *full-body interaction*—to explore in our exhibits, due to prior work indicating that they can make abstract concepts more concrete, approachable, and engaging [27, 35, 43, 45, 46, 54, 69]. Various forms of embodiment are also particularly effective at fostering learning in museums by making learning experiences apprehensible and enjoyable [7, 16]. We review prior work on the three forms of embodiment we focus on below.

Tangible user interfaces leverage physical materials as a way of interacting with digital interfaces [50]. Tangible user interfaces can make abstract concepts more concrete for learners when the physical materials invite learners to reflect on the materials' underlying representations [50]. This form of embodiment has proven effective at fostering learning about knowledge representations and machine learning in an at-home family learning activity, in particular by reducing intimidation and making AI more approachable [45]. Similarly, other work has found that tangible user interfaces leveraging familiar materials such as light can reduce intimidation in learning about AI [69].

"Stepping into the shoes" of an agent allows learners to use their own understanding of their body (i.e., their *body knowledge*) to reason about abstract concepts more concretely [8, 11, 23, 52, 63]. Learners can draw on conceptual metaphors [36] (i.e., reasoning about one idea through understanding of another) and embodied schemata [32] (i.e., structures of reasoning formed through bodily interactions) to make meaning. Prior work in both formal and informal learning contexts has demonstrated that this approach is successful in teaching non-technical learners about abstract AI concepts [22, 27, 46, 61]. For example, one study found that supporting learners' use of their body knowledge in the design of an exhibit about computational music remixing facilitated their sense-making and reasoning about interactions and patterns [46].

Full-body interaction allows learners to manipulate interface outputs using their body movements [30, 38, 49, 60]. This type of interaction has been shown to make learning about abstract concepts more intuitive by leveraging learners' unconscious understanding of movement and the relationships between movements [9]. Full-body interaction has been used to design learning interventions in mathematics [5], physics [49], and computing [33, 54]. In one study, learners at a remote camp run by a community organization connected abstract computing concepts to their concrete physical movements by creating computational artwork using dance movements as an input medium [54].

Although embodiment has been shown to support learning about AI and other abstract concepts in various informal learning environments, more research is needed to investigate what modes of embodiment are effective in promoting learning about AI in a museum setting.

2.3 Creativity & Learning

We consider creative activities to include ideation and improvisation, engagement with artistic practices, and generation of personally meaningful artifacts [14]. In various STEM learning environments, creativity is motivating; for example, creative activities such as computational music remixing [48] and e-textile construction [15] have fostered young learners' interest in computing. In more

recent work, creative activities (e.g., collaboratively drawing with a robot [66] or coding a game with an 'AI friend' [20]) and personally relevant activities (e.g., using generative AI to create a movie script about favorite video game characters [17]) have fostered young learners' interest in and engagement with AI. Additionally, creativity has been shown to facilitate personally relevant meaning-making, which can foster prolonged engagement (i.e., spending extended time interacting actively and meaningfully with exhibits) in a museum environment [3, 12, 24]. However, there have been no empirical studies examining how creative and personally relevant activities can foster learners' engagement with learning about AI in a museum setting. More research is needed to identify whether and how creativity can mediate young learners' engagement with an AI museum exhibit.

3 Exhibits Overview

We provide an overview of the exhibits we developed using a design-based research approach, which emphasizes iterative design and in-situ evaluation of learning experiences [1, 59]. As part of our design-based research process, we created conjecture maps [58] for each exhibit. Conjecture maps are a method commonly used in the learning sciences for simultaneously evaluating designs and theory in in-situ testing environments. Previous work has demonstrated the value of conjecture mapping for evaluating the design of informal learning experiences [51, 72]; thus, we leveraged the method to study our exhibits in a museum context. Conjecture maps hypothesize that certain *design* elements lead to *mediating processes* which produce the desired *outcomes*, and these three elements "provide the structure for mapping specific, testable conjectures of the relations among them" [58]. Our conjecture maps guide our design process, articulate how we hypothesize learning takes place in each exhibit, highlight the instantiation of the conjecture in our designs, and outline the observable processes (i.e., *mediating processes*) that lead to our desired learning and interest development outcomes.

We defined two overarching conjectures – one relates to *embodiment* and the other to *creativity* (Figure 2). From those, we developed exhibit specific conjectures that guided design and evaluation. We designed these conjecture maps by developing initial hypotheses about how embodiment and creativity may affect learning about AI based on existing literature. We then iteratively refined the conjectures throughout our design process and pilot tests. We present our final conjecture maps alongside the exhibit designs and exhibit learning objectives (which are based on existing guidelines for teaching young people about AI, including Long & Magerko's AI literacy competency framework [41] and the AI4K12 group's grade-level objectives for AI education [62]) below.

3.1 DataBites

3.1.1 Exhibit Conjectures. In our *embodiment* conjecture for *DataBites* (Figure 4), we hypothesize that allowing learners to physically manipulate data makes the abstract concept of a dataset more concrete and engaging, leading to better understanding of our first and second learning objectives (see Table 1). In our *creativity* conjecture, we posit that prolonged, open-ended, and fun creative engagement enables learners to connect familiar interests (such as their favorite pizzas and sandwiches) with AI, fostering sustained interest in AI

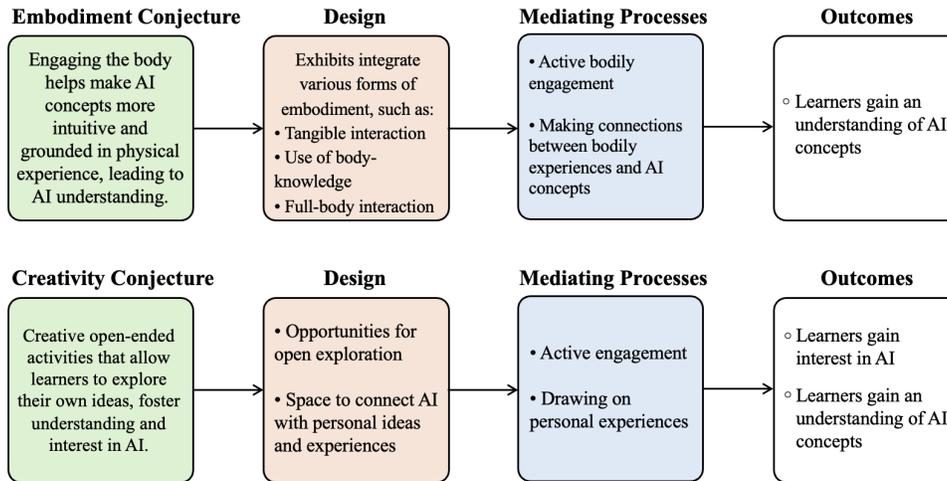


Figure 2: Overarching Conjectures: *Embodiment* and *Creativity*

Table 1: *DataBites* learning objectives. We refer to these learning objectives as DBLO# (DataBites Learning Objective #).

DataBites Learning Objectives
DBLO1: ML is a type of AI in which the computer learns patterns from many labeled examples.
DBLO2: Building a ML application involves multiple steps, including dataset preparation, training, and testing.
DBLO3: Human programmers decide what to include in the datasets of examples that are used to train AI.
DB LO4: The types of examples included in a training dataset affect the decisions the AI makes.



Figure 3: *DataBites* exhibit prototype. Top left: Tangible pieces for creating the training dataset. Bottom left: Buttons to start, train, and restart. Right: Full exhibit setup.

and a deeper understanding of the human role and the types of examples included in datasets when teaching AI (DBLO3, DBLO4).

3.1.2 Exhibit Design. The design of *DataBites*, informed by our conjectures, allows learners to train a computer to classify pizzas

and sandwiches by constructing a dataset with tangible pieces of pizza toppings and sandwich fixings (Figure 3). After constructing their dataset, learners hit the “Train” button to train an image classification pipeline. The pipeline learns patterns from an image of the learner-created dataset (captured using an overhead camera), then shows the classification results of a predetermined test dataset along with feedback to the learner via a wall projection (Figure 5). The machine learning pipeline is implemented as a CNN binary classification model, which leverages transfer learning to make final predictions highly sensitive to user input. The six learner-provided images are passed through the pretrained CNN to extract intermediate feature representations, which are then used to train a separate classifier exclusively on these images. This setup helps learners see how their dataset impacts the classification of the test images.

3.2 Knowledge Net

Table 2: *Knowledge Net* learning objectives. We refer to these learning objectives as KNLO# (Knowledge Net Learning Objective #).

Knowledge Net Learning Objectives
KNLO1: AI uses organized formats called knowledge representations to store information about ideas and how they connect.
KNLO2: AI is not all-knowing; its knowledge is restricted to the information people give it.
KNLO3: When you input incorrect or unusual information into an AI, it will generate incorrect or unusual outputs.

3.2.1 Exhibit Conjectures. The *embodiment* conjecture for *Knowledge Net* (Figure 6) theorizes that learners can utilize their body knowledge (i.e., their understanding of anthropomorphic bodies and motion) to reason about how their character will look and move

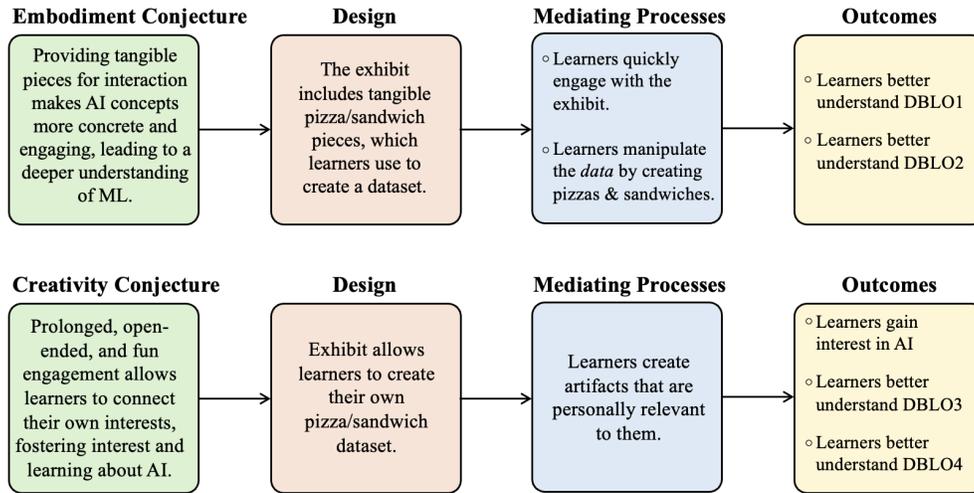


Figure 4: DataBites Conjectures: Embodiment and Creativity

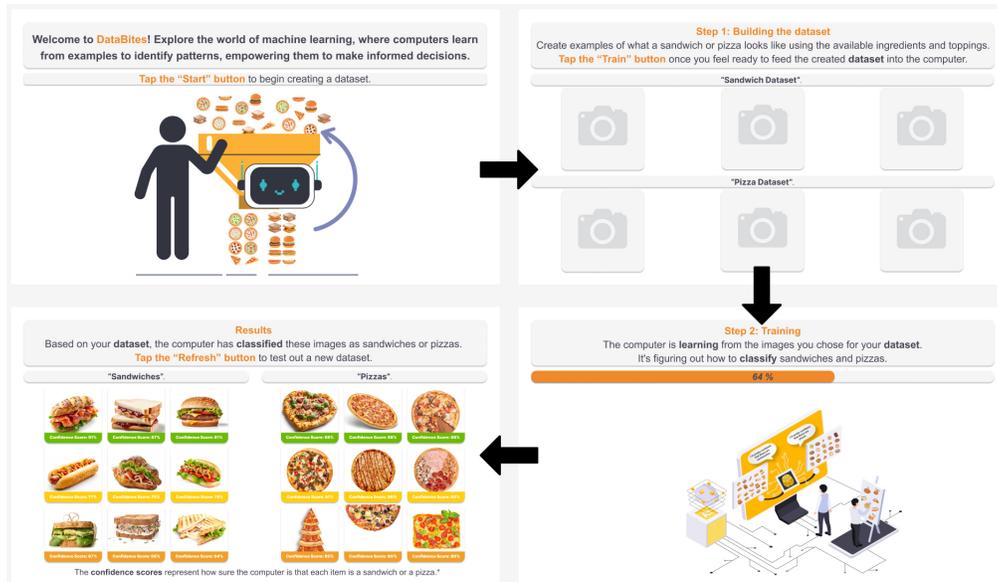


Figure 5: DataBites Exhibit UI. Top left: Start page with a brief intro to supervised ML, Top right: Live dataset creation, Bottom right: Animation of the training process, Bottom left: Classification results with confidence scores.

based on their inputs, leading to better understanding of our learning objectives (see Table 2). The *creativity* conjecture for *Knowledge Net* (Figure 6) posits that the familiar task of character personalization helps learners make associations between their inputs and the AI output, fostering understanding of that AI knowledge is restricted to the input information (KNLO2) and that inputting unusual information to a knowledge representation yields unusual outputs (KNLO3).

3.2.2 *Exhibit Design.* The *Knowledge Net* design, informed by our conjectures, features a terrarium-themed world in which learners

can design and control characters by building semantic networks representing the characters (Figure 7). The exhibit, built on an Ideum touch table using the Godot game engine, includes multiple stations for character creation to facilitate collaboration while still giving learners space for their individual creations [3]. Learners build a network through a drag-and-drop touch interface to control features of their character such as their hobbies, relationships to other characters, and physical attributes such as body color, shoes, and accessories (Figure 8). For example, the learner can create a network with the connections “Filbert IS blue” and “Filbert HAS

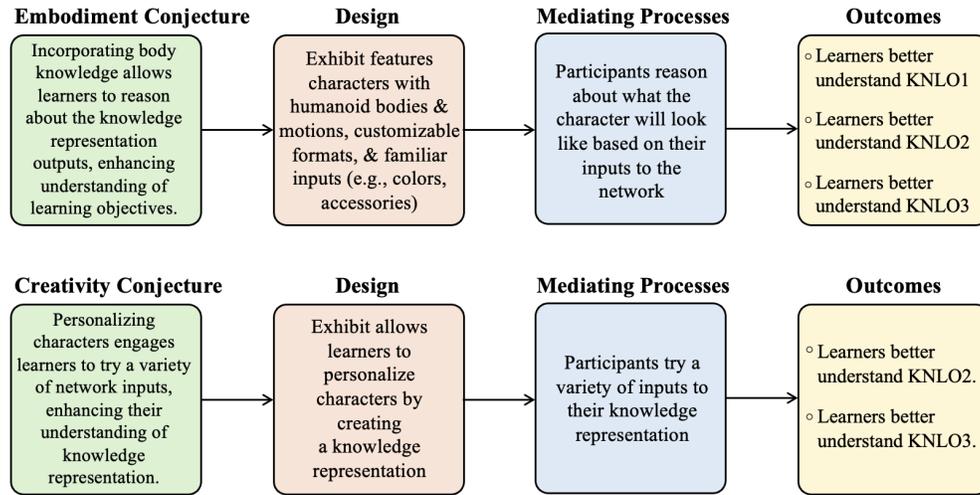


Figure 6: Knowledge Net Conjectures: Embodiment and Creativity



Figure 7: Top left: Close-up of Knowledge Net interface. Bottom left: Full view of four stations at Ideum table. Right: Participants interacting with Knowledge Net exhibit prototype.



Figure 8: Knowledge Net interface

book,” and the Filbert character would update in real-time to appear blue and hold a book. Characters also behave according to their semantic network by displaying a heart when they are near characters or qualities they LIKE and displaying a broken heart when near characters or qualities they DISLIKE. Learners can input conflicting facts, such as: Filbert LIKES Bliff and Filbert DISLIKES Bliff. In this case, Filbert would display both a heart and a broken heart when near Bliff to demonstrate the conflicting inputs.

3.3 LuminAIx

Table 3: LuminAIx learning objectives. We refer to these learning objectives as LAxLO# (LuminAIx Learning Objective #).

LuminAIx Learning Objectives
LAxLO1: AI can learn movement from a person.
LAxLO2: AI makes a decision based on learned data.



Figure 9: LuminAIx exhibit prototype.

3.3.1 Exhibit Conjectures. In our embodiment conjecture for LuminAIx (Figure 10), we hypothesize that full-body interaction through dance with an AI agent allows learners to make connections between their input dance moves and the agent’s output dance moves, leading to a better understanding of the learning objectives (see Table 3). In our creativity conjecture for LuminAIx (Figure 10), we

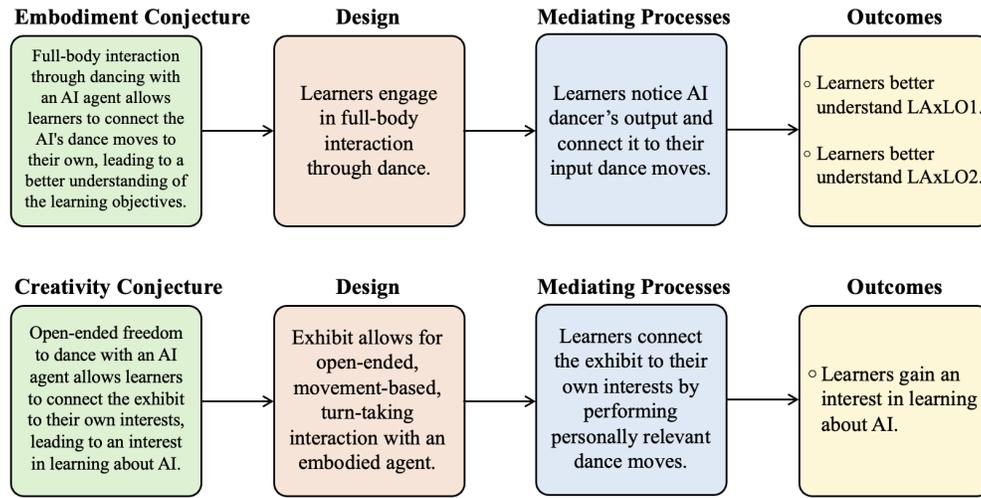


Figure 10: *LuminAIx* Conjectures: *Embodiment* and *Creativity*

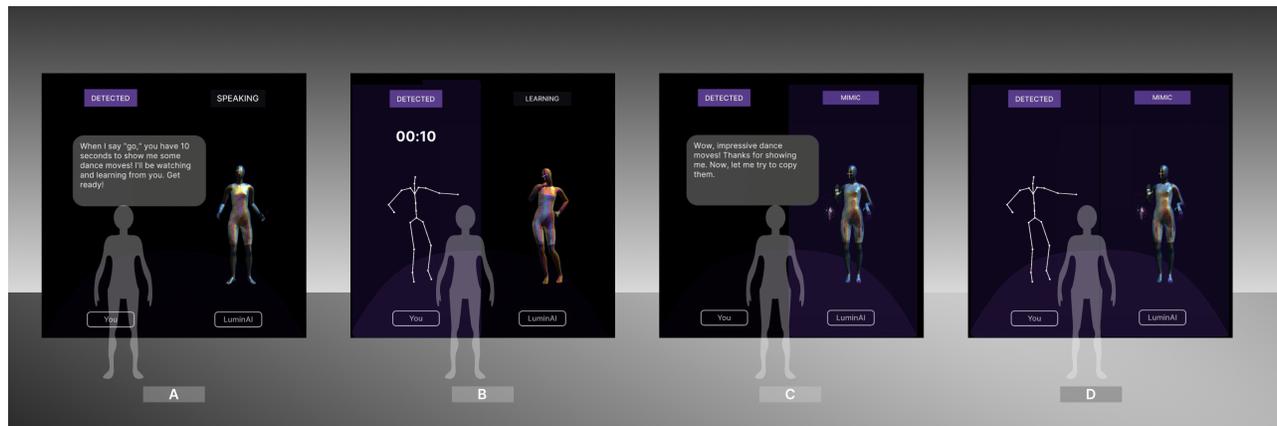


Figure 11: *LuminAIx* exhibit key interfaces.

hypothesize that open-ended freedom to dance with an AI agent allows learners to connect the exhibit to their own interests, leading to interest in learning about AI.

3.3.2 *Exhibit Design.* In *LuminAIx*, learners improvise dance movements with an AI dance partner. The installation is comprised of a screen projection (Figure 9) with a Kinect ¹ facing the learner to capture their live dance movements. Before the interaction begins, the screen displays *LuminAI*, the AI agent, swaying its body with a prompt “Step forward” to draw learners into the interaction zone. When the Kinect detects a learner steps in, the learner sees their skeletal representation mirrored on the panel and the interaction begins with the introduction remarks “Hello there, welcome to the dance floor! I’m *LuminAI*, your AI-powered dance partner. Can you teach me some moves?” Learners first engage in a guided dance interaction which is intended to help them learn about how

LuminAIx learns, responds, and collaborates. It starts with *LuminAI* saying “When I say ‘go’, you have 10 seconds to show me some dance moves! I’ll be watching and learning from you. Get ready!” after which the learner performs a dance move (Figure 11 B). *LuminAI* observes and responds by mimicking the participant’s movement (Figure 11 C). After the first trial, *LuminAI* then encourages the learner to teach it a more complex dance move. In another ten-second timed session, *LuminAI* generates an original response using its AI capabilities, creating movements contrasting with the learner’s moves. The user interface layout closely resembles that presented in Figure 11B-C. This scaffolded user flow aims to progressively familiarize participants with the interactive capabilities of *LuminAI* before fully immersing them in the open-ended dance experience.

Finally, after these two tutorial rounds, it transitions into an open-ended dance party. The participant and *LuminAI* engage in freestyle dance, collaboratively exploring and exchanging movements in

¹Microsoft Azure Kinect <https://azure.microsoft.com/en-us/products/kinect-dk>

real-time (Figure 11 D) that is not timed, and only finishes until the learner steps out of the interaction zone.

4 Methods

4.1 Data Collection

We conducted a study of the exhibits at a large science museum in the midwestern United States. We installed the exhibits for two consecutive days and recruited participants directly from the museum floor. This allowed us to study visitors engaging with the exhibits in-situ, in line with our design-based research methodology. We recorded visitor interactions with exhibits and aimed to recruit 30 participants to engage in interviews after their interaction. This sample size is in line with similar studies held on the museum floor [29, 42, 46].

Two cameras were set up to record visitor interactions with each exhibit: one to capture the participants' interactions with the exhibit and one to capture the exhibit interface. Signs indicated that cameras were recording and provided information about the research. To protect participants' privacy, all faces are blurred in published images. These video recordings are not the focus of our analysis in this paper.

We recruited participants for interviews by distributing flyers to families with children who appeared to fall within the target age range of 10–14 years, both on the museum floor and at the exhibits. After agreeing to participate, both the child and a parent were asked to provide written consent and age-appropriate assent, respectively, for participation in the interviews. Children who interacted with the exhibits and participated in interviews received \$10 Amazon gift card per exhibit that they answered questions about.

Five researchers conducted and audio-recorded interviews using a structured protocol (see supplemental materials). Participants who interacted with more than one exhibit were invited, but not required, to answer questions about each exhibit they visited; thus, participants participated in 1-3 interviews. On the first day, participants were interviewed about each exhibit in the same order of the study (i.e., *DataBites*, *Knowledge Net*, *LuminAlx*). On the second day, to mitigate ordering effects, we altered the order for each participant.

We developed an interview protocol rather than using an existing instrument because there are not many instruments for assessing AI understanding in our target age group [65], and the ones that do exist are not designed for museums or our specific learning objectives. We followed existing guidelines for interviewing children when designing the interview questions and duration [55]. We first asked about participants' prior knowledge about and interest in AI, followed by questions assessing how the exhibits fostered AI interest development. Finally, we asked content knowledge questions (Tables 4, 5, and 6) related to the AI concepts introduced in the exhibits. We provided worksheets and images as visual references for several questions to make the interview more child-friendly. This post-interaction approach to assessment was appropriate for an in-situ museum study, as pre/post evaluations interrupt the natural flow of the interaction on the museum floor [25]. Our full interview protocol and worksheets can be found in our Supplemental Materials. Interview questions are also presented in Tables 4, 5, and 3.

We classified our content knowledge questions according to Bloom's taxonomy [10], which describes different levels of cognition, to ensure that question difficulty was similar across all three exhibits. Most questions assessed the first (Remembering) or second (Understanding) level of Bloom's taxonomy, due to the brevity of the in-museum interactions. Because it is difficult to foster deep conceptual understanding in a museum setting, we instead aimed to measure high-level understanding of our learning objectives for each exhibit.

After the first day, we made minor adjustments to the interview to reduce the length due to participant fatigue. Specifically, we removed several retrospective questions that asked participants to reflect on how their interest in and knowledge of AI changed before and after interacting with the exhibits. We had used a similar retrospective question in a previous study where children engaged with tangible learning activities in their homes [45] so did not anticipate issues, but families participating in interviews on the museum floor have competing demands on their time (e.g., other exhibits to get to) and may be more easily fatigued after a long day at the museum. On the first day, interviews were running too long, causing parent frustration and child fatigue. We noticed that kids were speeding through the lengthy retrospective questions in particular without appearing to engage in reflection, so we made the decision to remove them.

We replaced these with three shorter open-ended interest-related questions, inviting participants to share what they liked or disliked about the exhibits and what initially attracted them. These changes provided participants with greater freedom to express their thoughts and reduced the length of the interview sessions. We do not report on the results from the retrospective questions asked on the first day as our observations indicated that kids were not engaging with them thoughtfully. This change resulted in some missing data about participants' interest from day 1. Our study received institutional IRB approval.

4.2 Data Analysis

4.2.1 AI Concepts and Interest Development. To assess whether learners understood the exhibit's learning objectives, we developed rubrics to analyze participants' responses to the content knowledge questions (see supplementary materials for full rubrics). We scored each interview question on a 4-point scale: 0 (inadequate answer), 1 (partial answer), 2 (adequate answer), and 3 (excellent answer). We engaged an evaluation expert to review and provide feedback on the rubrics.

We used an online platform (Rev.com) to obtain human-generated transcripts of the interview recordings, and we reviewed any sections that were inaudible or unclear. Two researchers then independently scored all responses and compared their assessments. Discrepancies were resolved through discussion.

To measure participants' maintained situational interest, an early stage of interest development [28], we quantitatively analyzed their responses to interest development questions (Yes/Maybe/No). We also calculated hold time (i.e., time participants spent interacting with the exhibit) from the video recordings as a marker of situational interest.

4.2.2 Creativity and Embodiment. To assess how creativity and embodiment affected participants' learning, we defined codes for these terms and coded for them in participants' responses to interview questions. We did not specifically ask participants about these concepts so that we could identify what aspects of creativity and embodiment in the exhibits were salient to them without priming them. We then analyzed how these factors influenced participants' learning and interest development in AI by comparing the codes they mentioned to their scores on the content questions and their self-reported level of interest in AI.

Based on existing literature, we defined several codes for creativity and embodiment. For creativity, we coded for:

- fluency (trying a variety of inputs) [19]
- making (creating an artifact) [19]
- uniqueness (trying out improbable inputs) [14]
- personal relevance (using familiar inputs) [14]

For embodiment, we coded for:

- tangible interaction: physical manipulation of the exhibit (*DataBites*) [50]
- full-body interaction: physical dance moves (*LuminAlx*) [30, 38, 49, 60]
- body knowledge: character bodies and motion (*Knowledge Net*) [8, 11, 23, 52, 63]

Three researchers coded participants' responses as a group and resolved any conflicts through discussion. During this first round of coding, we refined our code definitions based on how participants talked about their experiences and we noted exhibit-specific examples of each code. We also inductively added two new codes, anthropomorphism (relating AI to humans) and mechanization (attributing AI to its hardware components), as we encountered these themes in participants' responses and found them relevant to our research questions. The three researchers then re-coded all of the responses based on our updated codebook. See supplementary material for our full codebook with exhibit-specific definitions. In addition to coding interview responses, we also looked to the video recordings to confirm that participants were interacting with the exhibits in the ways we expected (e.g., manipulating the tangible materials).

5 Results

In the following sections, we evaluate each exhibit's conjectures by synthesizing results from our analysis of participants' understanding of AI concepts, qualitative coding of interviews, video observations, and reported interest data. For each conjecture, we use our data to verify the salience of the relevant *design* element and check whether the *mediating processes* and intended *outcomes* occurred.

5.1 Demographics

We conducted interviews with 29 10-14 year olds. We interviewed 24 of these participants about *DataBites*, 22 about *Knowledge Net*, and 18 about *LuminAlx*. Of our 29 participants, 44% (n=13) were female, 34% (n=10) were male, 10% (n=3) were non-binary, and 10% (n=3) did not report gender identity. 44% (n=13) were White, 13% (n=4) were Hispanic or Latinx, 10% (n=3) were Asian, 3% (n=1) were

Black or African-American, 10% (n=3) were multi-racial, and 17% (n=5) did not report their race/ethnicity. Parents reported that 31% (n=9) participants had very little prior knowledge about AI, 48% (n=14) had some prior knowledge about AI, 10% (n=3) had a lot of prior knowledge about AI, and 10% (n=3) did not report on prior knowledge about AI.

5.2 Embodiment Conjectures

5.2.1 DataBites. To evaluate the *DataBites* embodiment conjecture (see Figure 4), we analyze the salience of tangible manipulation of data for participants (*design*), whether participants were highly engaged with the tangible components of the exhibit and demonstrated that they were tangibly manipulating data (*mediating processes*), and whether participants demonstrated understanding of the first two learning objectives (*outcomes*).

Design: Did learners find tangible interaction to be a salient part of their experience? 10 of 24 participants (41.7%) mentioned physically creating the pizza and sandwich dataset items and engaging in the hands-on process as a salient feature of *DataBites*. Moreover, when asked what they liked about the exhibit, several participants noted these interactive aspects (P15b: "It looked fun, moving all of the pieces around and then seeing the end.")

Mediating Processes: Did learners demonstrate high engagement with the tangible components of DataBites? Our analysis of mediating processes revealed that learners were highly engaged with the tangible components of the *DataBites* exhibit. Video recordings confirmed that all participants interacted with the tangible elements, even if they did not explicitly mention them in their interviews. Furthermore, the videos showed that participants spent more time engaging with the dataset creation phase (over 3 minutes), which was tangible and interactive, compared to the testing phase (under 50 seconds), which was mostly passive and involved only observing results. To ensure that the engagement with tangibles aligned with the exhibit's intended learning goal—creating a dataset—we included a question in the interviews asking participants to either explain or recreate what they had built. They were given a worksheet that mirrored the dataset structure from the exhibit, along with tangible pizza and sandwich pieces. We then analyzed their responses to determine whether they were meaningfully constructing datasets or simply placing random items. Most participants (n=16, 66.7%) chose to draw or recreate their dataset using the tangible pieces, while the rest described it verbally. All participants described or constructed at least one data point that was relevant and coherent within the pizzas or sandwiches dataset (e.g., combinations of ingredients plausibly representing a pizza or sandwich). None created entirely random or unrelated combinations. This shows that they were actually engaged with the pizzas and sandwiches as data rather than randomly manipulating the pieces.

Outcome: Did learners demonstrate understanding of DBLO1 and DBLO2? To assess DBLO1, we asked participants to identify how the pizza/sandwich examples they provided affected the AI (Table 4). In their responses, most participants (n=17, 70.8%) acknowledged that the training dataset had an effect on the AI's understanding of pizzas and sandwiches, but only a few (n=4, 16.7%) recognized that

Table 4: DataBites interview questions and results.

Interview Question	Mean Score	Standard Deviation	Interquartile Range	Learning Objective	Bloom's Taxonomy Level
Q1: Can you describe some of the steps involved in teaching a computer to recognize pizzas vs. sandwiches?	0.71	0.69	1	DBLO2	Remembering
Q2: How did your examples affect the AI's ability to recognize pizzas and sandwiches?	0.92	0.78	1	DBLO1, DBLO3	Understanding
Q3: What do you think would happen if you include only examples of one type of pizza or sandwich in your training dataset?	1.46	0.83	1	DBLO4	Understanding
Q4: What might happen if you put pizzas in your sandwich training dataset and put sandwiches in your pizza training dataset?	1.42	0.78	1	DBLO4	Understanding
Q5: With DataBites, you were trying to train the AI to recognize pizzas and sandwiches. What would happen if you trained an AI to tell the differences between the faces of your friends and family? Could you see any issues or challenges with that?	0.88	0.74	1	DBLO4	Applying

the AI learns from patterns among this dataset. Thus, the exhibit was fairly successful at demonstrating that AI learns from labeled data, but needs further refinement to establish the importance of patterns among that data.

To assess DBLO2, we asked participants to identify the steps of ML. About half of the participants ($n=14$, 58.3%) identified at least one step of the process. Most of these participants ($n=11$) identified the dataset creation step, which was the focus of the pizza/sandwich creation portion of the activity. Only a few participants ($n=3$, 12.5%) named all three steps. This suggests that the exhibit was notably more successful at introducing the dataset preparation step than teaching about the other steps, although the salience of all steps could be improved in future iterations.

Conjecture: Did finding tangible interaction to be a salient feature of the exhibit correspond with higher learning scores? 10 of 24 participants (41.7%) mentioned tangible interaction during their interviews. Eight of the individuals who mentioned tangible interaction ($n=8$, 80%) scored above average on the question assessing DBLO1, and six of those who mentioned tangible interaction (60%) scored above average on the question assessing DBLO2. These results suggest that there may be a relationship between tangible interaction and the participants' understanding of the learning objectives. This in-situ finding could be further substantiated in a more controlled study environment.

When we asked about the steps involved in building a machine learning application (DBLO2), we anticipated responses covering dataset preparation, training, and testing. However, most participants ($n=17$, 70.83%) only mentioned the dataset preparation phase, which was the only phase that was conveyed via tangible interaction. This suggests that participants understood concepts better when they were presented through tangible interactions. This highlights the importance of incorporating more interactive elements into the training and testing phases.

Based on these results, our conjecture—which theorized that tangible interaction would engage participants in manipulating data and lead to learning (Figure 4, bottom)—was supported. However, further refinement is needed to ensure that tangible interactivity remains consistent throughout all phases of the exhibit, not just during dataset creation. Future design iterations could also more clearly highlight the role of patterns in the dataset to make them more salient to learners.

5.2.2 Knowledge Net. To evaluate the *Knowledge Net* embodiment conjecture (Figure 6), we analyze the salience of the humanoid bodies and motion in the experience for learners (*design*), whether participants reasoned about the outputs of the knowledge representations (*mediating processes*), and whether participants understood the *Knowledge Net* learning objectives (*outcomes*).

Table 5: Knowledge Net interview questions and results.

Interview Question	Mean Score	Standard Deviation	Interquartile Range	Learning Objective	Bloom's Taxonomy Level
Q1a: Imagine you create a network that looks like the network below (Figure 12, left). Please describe or draw what the output character might look like.	1.32	0.57	1	KNLO1	Understanding
Q1b: Given this network (Figure 12, left), what is NOT a possible feature of the character?	1.64	0.79	0	KNLO1	Understanding
Q2: Looking at the image below (Figure 12, right), could you draw a network that represents these characters?	1.14	0.85	0	KNLO1	Understanding
Q3a: Please draw a network to teach an AI about yourself. Imagine that the AI would use your network to generate a visual of you and answer questions about you if asked.	1.38	1.07	2	KNLO1	Applying
Q3b: What do you think the AI would do if you put false information about yourself in this network?	1.48	1.12	2	KNLO2, KNLO3	Understanding
Q3c: Is there anything this network does not capture about you?	1.38	1.12	2	KNLO2	Understanding

Design: Were the humanoid character bodies and motions salient to learners? In the interview responses, only a few participants (n=3, 13.6%) explicitly mentioned the embodied aspects of the characters, such as their movements or interactions (P14b: “I liked, based on your choices, your character would go somewhere, like if the same characters liked each other and they had basically the same, they would go closer to each other”). This indicates that the familiar character bodies were not necessarily a salient part of the interaction for most participants.

Mediating Process: Did learners reason about the outputs of their networks? In our video analysis, we found that no participants verbally reasoned about how their characters would change based on their knowledge representation inputs. The exhibit featured signage asking, “What happens if both characters like each other?”, and few participants (n=2, 9.1%) engaged with this signage by reading it aloud and verbally saying, “let’s figure it out.” However, these participants simply tested out what would happen by manipulating their knowledge representation rather than verbally hypothesizing. Thus, participants did not explicitly utilize their body knowledge to reason about the outputs of their networks.

Outcome: Did learners demonstrate understanding of KNLO1, KNLO2, and KNLO3? To assess KNLO1, we asked several questions (see Table 5). First, participants were shown a picture of a semantic network similar to those they may have constructed while interacting with

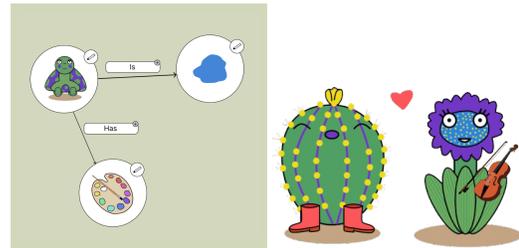


Figure 12: Left: Knowledge Net Q1 network. Right: Knowledge Net Q2.

the exhibit and were asked to draw or describe the corresponding character. Only some participants (n=7, 31.8%) correctly included all of the features in the semantic network, but more participants (n=12, 54.5%) included some, though not all, of the correct features. Then, we asked participants which of three facts could not be true about the character represented by the network (Figure 12). Most participants (n=18, 81.8%) correctly chose that the character would not hold a book based on this representation. Then, participants were shown an image of two characters from the exhibit (Figure 12, right) and asked to draw the semantic network(s) that would generate these characters. Most (n=17, 77.3%) produced a semantic network, even if it was incorrect topically. Finally, participants were

asked to draw a network to teach an AI about themselves; most (n=16, 72.7%) were able to produce a semantic network representing some facts about their lives. Overall, these results indicate that while participants may not have grasped the details of semantic networks, they generally understood that a knowledge representation shows how ideas connect and that the knowledge representation influenced the character's appearance.

To assess KNLO2, we asked two questions (see Table 5). First, we asked what the AI would do if given incorrect information. Most participants (n=14, 63.6%) recognized that the AI output would reflect the given incorrect information. Then, we asked participants if there was anything the network they created did not capture about their lives. Some participants (n=10, 45.5%) identified at least one aspect of their lives that was missing from the network, indicating awareness of the limitations of knowledge representations. Additionally, three participants (n=3, 13.6%) explicitly stated that the AI could never fully know everything about them (P4: "No, you can't teach it everything about me, because there is so much to learn about myself"). These results suggest that only some participants understood that a knowledge representation is limited to the information provided to it by humans.

Finally, to assess KNLO3, we asked participants what the AI would do if given incorrect information in the network. More than half of the participants (n=14, 63.6%) responded that the AI's output would reflect the incorrect input. Thus, most participants understood that the output of a knowledge representation is limited based on its input.

Conjecture: Did utilizing body knowledge correspond with higher learning scores? In this section, we explore whether the use of body knowledge supported participants in their understanding of KNLO1, KNLO2, and KNLO3. We found that the humanoid bodies and motions of the characters in the exhibit were not salient to the participants. Additionally, we found no evidence that learners reasoned about the outputs of their knowledge representations. However, we found that participants at least partially understood our learning objectives.

Based on these results, we cannot conclude that the design of the exhibit facilitated the use of participants' body knowledge to understand the learning content. Further work is needed to investigate which mechanisms of the exhibit helped learners understand the learning objectives, whether learners were drawing on their prior knowledge and reasoning capabilities to answer the evaluation questions rather than their exhibit experience, and whether making body knowledge a more salient part of the exhibit design could lead to improved learning. Interestingly, we found that the highest scorers (P14b and P19b) both mentioned that they liked the embodied characters and their interactions, and P14b additionally mentioned that he enjoyed the character movements. This indicates the potential that the embodied nature of the characters can facilitate understanding of the learning content.

Based on these results, our conjecture that the use of familiar character bodies and inputs would lead to the prediction of outputs and facilitate learning was not supported. Future designs will focus on strengthening the connections between the network inputs and character outputs (e.g., only having characters move if learners give them that ability), fixing UI issues that may have hindered

observations of characters' behavior (occasionally the characters got 'stuck' when colliding with each other as they moved), and implementing more character interactions to make the characters' embodiment of the network more noticeable to learners. We also plan to iterate on our evaluation questions to ensure that they draw on learners' exhibit experience rather than any existing prior knowledge they may have. Once we successfully facilitate learners' use of body knowledge in the exhibit, we can test its role in learning about knowledge representations.

5.2.3 LuminAIx. To assess the *LuminAIx* embodiment conjecture (Figure 10), we look at the salience of full-body interaction in the exhibit (*design*), whether the participants notice the AI output and connect it to their own moves (*mediating process*), and whether participants understood the learning objectives of *LuminAIx* (*outcome*).

Design: Was full-body interaction a salient part of the experience for learners? On day 2, when we asked our 12 participants what drew them to and/or what they liked about *LuminAIx*, a little over half (n=7, 58.3%) mentioned that they enjoyed engaging in full-body interaction (P12: "I liked the dancing. It was fun"), indicating that this aspect of the experience was salient for them.

Mediating Process: Did learners connect the AI agent's moves to their own? On day 2, when we asked our 12 participants what they liked about *LuminAIx*, most (n=8, 66.7%) mentioned that they enjoyed how the AI agent responded to their dance moves. Of these 8 participants, most (n=5) noticed that the AI agent repeated their dance moves (P18b: "I thought it was kind of cool how that AI person copied what you did and how it sensed it"), whereas only a few (n=3) noticed that the AI performed similar, though not the same, moves (P18a: "[I liked] how the AI dancer mixed up the movements that you did").

Inductively, we observed that participants conceptualized the AI dance partner using two different mechanisms. Some participants (n=4, 33.33%) used only *anthropomorphism*, attributing human traits to AI (P2b: "Maybe it just has the mind of its own and it does its own dance moves"). Others (n=5, 41.67%) used *mechanization*, reducing AI to its hardware and technical components (P22: "It tracks your body and then I guess it just puts the same coding into the thing and just makes it look like more human"). Additionally, some participants (n=8, 44.44%) used both anthropomorphism and mechanization. This indicates that it was salient to participants that the AI agent reacted to their moves, and they conceptualized the AI agent as either a human-like entity or a more technical artifact.

Outcome: Did learners demonstrate understanding of LAxLO1 and LAxLO2? We asked two questions to assess LAxLO1 (Table 6). First, we asked participants to hypothesize how the AI agent knew its dance moves. In their responses, most participants (n=16, 88.9%) demonstrated understanding that the AI responded to their own movements. Then, we asked participants how the agent's moves were similar or different from their own moves and why they believed those similarities or differences occurred. Most participants (n=14, 77%) were able to identify either similarities or differences between their moves and the AI's, but provided vague (or no) explanation of why the moves were similar or different. Based on these

Table 6: *LuminAIx* interview questions and results.

Interview Question	Mean Score	Standard Deviation	Interquartile Range	Learning Objective	Bloom's Taxonomy Level
Q1: Remember when you were dancing with the AI partner. How do you think the AI agent knew those dance moves? Why?	2.11	0.93	1	LxLO1	Understanding
Q2: Can you describe how different or similar your dance movements with the AI partner were? Why did you think this happened?	2.11	0.58	0	LxLO1	Analyzing
Question 3: Remember when you had a dance party with an AI partner. How do you think the AI partner chooses which dance moves to do while dancing? Why?	1.44	1.15	2	LxLO2	Analyzing
Question 4: At the dance party, what do you think would happen if you danced differently?	1.50	1.10	1	LxLO2	Evaluating

results, most participants seemed to grasp that the AI can learn movement from a person.

To assess LxLO2, we first asked participants how they thought the AI agent chose its movements during the dance party. Half of the participants ($n=9$, 50%) were able to identify that the AI not only observes their movements, but then uses that data to make a decision about what movement it performs. Then, we asked participants what they thought would happen if they danced differently during the dance party. Half of the participants ($n=9$, 50%) acknowledged that the AI agent would not only dance differently, but do so according to the user's new moves. These results indicate that participants had a limited understanding that the AI agent made decisions about its moves based on their own moves.

Overall, these results demonstrate that most participants understood that the AI agent learns from their dance moves, but fewer understood that their moves are used by the AI to make a decision about its own moves.

Conjecture: Did finding full-body interaction to be a salient feature of the exhibit correspond with higher learning scores? We found that a little over half of the participants ($n=7$) interviewed on day 2 reported enjoyment of their own full-body interaction through dance. It's possible that more participants enjoyed this feature, but as we added this question on day 2, the day 1 participants could not report it.

When we asked what they liked about *LuminAIx*, over half ($n=4$) of the participants who mentioned enjoying dancing additionally acknowledged that the digital agent not only danced, but danced in a manner responsive to their own moves. This indicates that while participants may have enjoyed dancing, the fact that they danced with an AI agent was only salient to some of them.

All four of the participants who enjoyed dancing with the AI agent demonstrated understanding in the interviews that the AI agent learned their movement (LxLO1), but only two of these four participants demonstrated full understanding that the AI makes decisions based on their moves (LxLO2). The other two participants only demonstrated partial understanding of this learning objective.

Regardless of whether they explicitly mentioned enjoying the dancing aspect of the exhibit, most participants ($n=8$) who were asked about what drew them to the exhibit and what they liked about the exhibit mentioned the fact that the AI agent responded to their dance moves in some way. This indicates that even if the participants did not explicitly state that they enjoyed dancing, it provided a clear input for the AI to act upon.

Overall, most participants demonstrated understanding of LxLO1, while they only partially understood LxLO2. Additionally, participants who used *both* mechanization and anthropomorphism when talking about the AI agent scored the highest on average across all questions ($n=8$, $M=2.03$) when compared to participants who only used mechanization ($n=5$, $M=1.8$), participants who only used anthropomorphism ($n=4$, $M=1.44$), or participants who used neither ($n=1$, $M=1.25$).

These results indicate that using full-body interaction through dance as the input and output of an AI system allows learners to make connections between those inputs and outputs; however, more work is needed to foster understanding of what the AI system does with the inputs and how it makes decisions. Learners who conceptualized AI as possessing a mixture of human and non-human traits understood our learning objectives the best.

Our conjecture that full-body interaction can act as an input/output type in which learners can connect their inputs and the AI outputs was supported, but more work (with a larger sample size) is needed to verify whether this interaction type mediates understanding of

our learning objectives. In the next design iteration, we plan to add additional text to the user interaction to explicitly state that the AI agent chooses its dance moves based on the learner's moves, include text-to-voice to make sure learners are not missing key information, and provide interactive visualizations of how the AI uses learned moves in its decision-making.

5.3 Creativity Conjectures

5.3.1 DataBites. To assess our *DataBites* creativity conjecture, we analyze whether participants were engaged in creating their own datasets of pizzas and sandwiches (*design*), created personally relevant artifacts and explored a variety of artifacts (*mediating processes*), and demonstrated interest in learning about AI and understanding of learning objectives DBLO3 and DBLO4 (*outcomes*).

Design: Did learners exhibit active and prolonged engagement? The average hold time for all visitors at the *DataBites* exhibit (not just those we interviewed) was 3.61 minutes (n=56), exceeding the 3.3-minute threshold typically associated with active prolonged engagement [3]. Additionally, our video recordings involving 12 participants showed that all 12 were actively engaged, and the open-ended nature of the task gave them the freedom to create their own datasets in different ways. Notably, six of them chose to complete the activity a second time, suggesting prolonged engagement and maintained situational interest [28].

Mediating Processes: Did learners create personally meaningful artifacts? When asked about what types of pizzas and sandwiches they made at the exhibit, 12 of 24 participants (50%) shared that they created pizzas and sandwiches they had personal connections to, such as pizzas or sandwiches they commonly ate with their families. Learners enjoyed connecting their tastes in food to their creations and incorporating their personal interests into the design of their pizzas and sandwiches (P20a: “[I liked] that you could teach AI what type of food you like, kind of”). They also shared that the items they created evoked memories and reflected their personal tastes (P5a: “I made a burger, which actually reminds me of what I got to eat with my parents”). When asked about their experience and impression of the exhibit, 13 of 24 participants (54.2%) expressed that they enjoyed the freedom to create imaginative items, especially ones that allowed them to try new ideas (P17b: “You can do funny faces. And you can put anything you want in the pizza or sandwich”). These results indicate that participants recognized creativity as an important part of their experience at *DataBites*, especially the ability to explore and incorporate their personal interests.

Outcome: Did learners demonstrate understanding of DBLO3 and DBLO4? To assess DBLO3, we asked participants whether the AI learned from the data they provided (Table 4). Most participants (n=17, 70.8%) demonstrated an understanding that AI learned from the data they provided. This suggests that most participants grasped the role humans play in training AI systems.

To assess DBLO4, we asked several questions (Table 4). Most participants (n = 20, 83.33%) recognized that AI is influenced by its training data. Among them, 14 participants (58.33%) understood that a limited dataset leads to limited results. The majority of participants (n = 21, 83.33%) showed awareness that AI is affected by its training data, and of those, 12 participants (n=12, 50.0%) recognized

that using incorrect data can lead to incorrect results. 16 participants (66.7%) recognized that the types of examples provided affect what the AI can do.

Outcome: Did learners demonstrate an interest in learning about AI? When we asked about how the exhibit sparked interest in AI, most found the exhibit enjoyable (n=22, 91.66%) and interesting (n=22, 91.66%), and expressed strong interest in learning about how AI works (n=18, 75%), and participating in similar activities in the future (n=19, 79.16%). Also, *DataBites* supported active prolonged engagement. This suggests the exhibit successfully fostered interest in learning about AI.

Conjecture: Did finding the open-ended task to be a salient feature of the exhibit correspond with higher learning scores and interest in AI? In the interviews, about half of the participants (n=12, 50%) explicitly and unpromptedly mentioned creating personally relevant datasets, indicating this aspect of the exhibit was salient to them. Others may have done so as well, but did not mention it in their interviews (as the question about what they liked about the exhibit was added on day 2 of the study). Of the 12 participants who mentioned personally meaningful creation, over half (n=7, 58.3%) expressed that they wanted to learn more about AI in the future, and some others (n=4, 33.3%) mentioned being open to learning more about AI. This suggests the exhibit successfully fostered situational interest in learning about AI. These results emphasize the success of the exhibit in maintaining situational interest and engagement, supporting the conjecture that open-ended, personally relevant experiences can spark and maintain curiosity about AI. Results from the analysis of learning outcomes suggest that most participants achieved DBLO3 and DBLO4, but participants who described creating personally relevant dataset items did not perform any better than other participants. This suggests that finding the personal relevance salient was not necessarily connected to a deeper understanding of DBLO3 and DBLO4. The conjecture that creative and prolonged engagement can maintain situational interest in AI and support participants in learning DBLO3 and DBLO4 (Figure 4, top) was partially supported: the exhibit fostered interest in AI and learning, but the connection between personal relevance and learning was unclear. To enhance learning in future iterations, the design will incorporate more elements aligned with the exhibit's learning objectives (e.g., additional AI-related visuals on the UI and physical components, audio cues, etc.). These additions aim to deepen understanding while preserving the exhibit's creative and open-ended nature.

5.3.2 Knowledge Net. To assess our *Knowledge Net* creativity conjecture (Figure 6), we analyze whether the personalization of characters engaged participants (*design*), whether participants tried a variety of inputs in their networks (*mediating process*), and whether participants understood the KNLO2 and KNLO3 learning objectives (*outcome*).

Design: Were learners engaged by the personalization aspect of the exhibit? On day 2, when we asked 13 participants about what drew them to *Knowledge Net* and what they liked about the exhibit, a little over half (n=8, 61%) mentioned that they enjoyed creating their own character (P20a: “It was interesting to me how you could

just customize whatever character you used”). This indicates that personalization was a salient feature of *Knowledge Net*.

Mediating Process: Did participants try a variety of network inputs? In our video analysis of interviewees' interactions with *Knowledge Net*, we observed that only some of our participants (n=10, 45.0%) constructed networks with a variety of inputs. Additionally, a few (n=2) participants mentioned that they enjoyed trying different inputs in their interview responses. Thus, the exhibit design did not successfully encourage participants to try out a variety of inputs in their knowledge representations.

Outcome: Did learners demonstrate an understanding of KNLO2 and KNLO3? In this conjecture, we focus on participants' understanding of KNLO2 and KNLO3. We found that most participants understood both learning objectives (see Section 5.2.2 for score breakdowns by learning objective).

Conjecture: Did finding personalization a salient feature of the exhibit correspond with higher learning scores? We found that personalization was moderately salient in *Knowledge Net* as 61% of participants (n=8) who were asked about their favorite aspects of the exhibit mentioned this feature. Not all participants who expressed enjoyment of the character personalization tried a variety of inputs, and vice versa. It is possible that some of the participants who tried a variety of inputs did, in fact, enjoy the personalization, but were not asked about what they liked about the exhibit (as this question was added on day 2 of the study). Of the eight participants who mentioned personalization in their interviews, we observed that only some (n=4) tried out a variety of inputs in their semantic networks. Three of these four participants demonstrated partial understanding of KNLO2, and all four of these participants demonstrated understanding of KNLO3. These results indicate that only some of the participants who were engaged by the character personalization feature were motivated to try a variety of inputs in their networks, but those who did understand KNLO2 and KNLO3.

We observed that 10 of 22 total participants (45.5%) tried a variety of inputs. Nine out of ten answered the interview questions relevant to these learning objectives as one participant ended the interview early. All of the nine participants who tried a variety of inputs and answered the relevant questions (n=9, 100%) demonstrated at least partial understanding of KNLO2, and most of these participants (n=8, 88.9%) demonstrated understanding of KNLO3. This indicates that participants who tried a variety of inputs understood KNLO2 and KNLO3.

Overall, these results demonstrate that personalization was a salient feature of *Knowledge Net*, but that participants who enjoyed the personalization feature were not necessarily motivated to try out a variety of inputs in their networks. Trying a variety of inputs seemed to lead to understanding of KNLO2 and KNLO3. This suggests that trying a variety of inputs is important to the learning goals of the exhibit.

Our results supported our conjecture that personalization can engage learners (Figure 6, top), but there is still room for improvement in motivating users to try a variety of inputs, which seems to foster better understanding of our learning objectives. We hypothesize that some participants may have been discouraged from trying many inputs due to a network creation interface that some

found challenging to use (e.g., the drag and drop interaction was not clearly intuitive), an incomplete input/output mapping (e.g., there was no corresponding output if the learner created a network with the connection “Filbert IS book”), and the aforementioned issue in which characters occasionally got ‘stuck’, limiting their interactions. In the next step of the design process, we will refine the interaction to improve the network creation UI, provide a wider variety of input options, ensure a complete input/output mapping, and increase interaction between the characters.

5.3.3 LuminAlx. To assess our *LuminAlx* creativity conjecture (Figure 10), we analyze whether participants found creative freedom a salient feature of the exhibit (*design*), whether they demonstrated exploration and personally relevant creativity in their interactions with *LuminAlx* (*mediating processes*), and if they demonstrated interest in learning more about AI (*outcome*).

Design: Was the open-ended dance activity salient for learners? When we asked participants on day 2 about their favorite aspects of *LuminAlx*, the majority of the 12 participants (n=10, 83.33%) expressed enjoyment at the opportunity to create a dance with the AI agent. This indicates that creating dance moves was a salient feature of *LuminAlx*.

Mediating Processes: Did learners perform personally meaningful dance moves? Two participants mentioned that engaging with the exhibit allowed them to connect with their own creativity (n=2, 16.66%). P14a remarked, “I liked that I saw what I did in the end,” reflecting on the satisfaction of seeing their movements captured and represented by the installation. One participant (n=1, 7%) felt personally connected to the exhibit as it related to their own interests (P20a: “There was music playing, and then it just felt good because I like music”).

Additionally, the exhibit encouraged exploration and experimentation, as video analysis revealed that four participants (n=4, 22.22%) tried out different dance movements (P21: “You got to show off your moves”). This sense of freedom to experiment fostered a playful and exploratory environment where participants could engage in improvisation, express their personalities, and enjoy the feedback loop between their creativity and the AI's responses. These findings indicate that personally relevant movement and exploration were featured in some participants' interactions with the exhibit.

Outcome: Did participants demonstrate interest in AI? When we asked interest development questions about the exhibit, most participants found it enjoyable (n=17, 94.44%) and interesting (n=15, 83.33%), with a large majority expressing interest in learning more about AI (n=14, 77.77%). Many reported that the activities encouraged deeper thinking about AI (n=10, 55.55%), and a large portion showed enthusiasm for doing similar activities in the future (n=13, 72.22%). These results suggest that the exhibit was successful in engaging participants and fostering sustained interest in AI learning. Additionally, we looked at participants' responses to identify further evidence of interest development. Some participants were drawn to the motion tracking and real-time responsiveness of the AI system (P18a: “How it did well at tracking the movements without any delay”). Others expressed curiosity about the AI's ability to mirror and interpret human movement (P18b: “I thought it was kind of cool how that AI person copied what you did and how it sensed it”).

The AI’s interactive behavior also introduced a sense of challenge and playful competition (P19: “I thought it was an interesting idea that the AI was going to try and copy what you did, and I took it as kind of a challenge to beat it, I guess”). These observations highlight how the exhibit’s design features—particularly its responsiveness and mimetic capabilities—fostered participants’ engagement and curiosity about the technology underlying LuminAI.

Conjecture: Did finding the open-ended dance activity a salient feature of the exhibit connect to interest in AI? In this section, we investigate how personally relevant movement and exploration sparked participants’ interest in AI. Data from participants’ interactions show that most participants enjoyed creating different and personally meaningful dance moves with the AI agent. At the same time, analysis of their interest development indicates that the exhibit successfully sparked interest in AI for the majority of participants. As these two indicators were consistent across most participants, we were not able to directly map them onto one another. Nonetheless, their simultaneous presence provides suggestive evidence that these features of the exhibit contributed to participants’ interest development.

Our data provides partial evidence supporting the conjecture that personally relevant movement and exploration contributed to participants’ interest in AI. Since we could not clearly assess whether participants’ dance moves were personally meaningful or directly linked to their interest in AI, future work will include interview questions about the motivation behind their movements to explore how meaningful or exploratory actions contributed to sparking interest in AI.

6 Discussion

In this paper, we aimed to answer two research questions: RQ1) *How can we leverage embodiment and creativity to design museum exhibits to foster middle schoolers’ AI literacy?*; and RQ2) *How do embodiment and creativity influence the development of middle schoolers’ AI literacy and interest in AI?*

In response to RQ1, we presented the designs of *DataBites*, *Knowledge Net*, and *LuminAIx*, three novel exhibits that use embodiment and creativity to teach learners about AI. Our iterative design process allowed us to explore different kinds of embodiment and creativity across these exhibits and to refine how these elements supported AI learning. Through iterative cycles, we examined how tangible, bodily, and full-body forms of interaction—as well as opportunities for free-choice creation and personalization—can be effectively integrated into AI learning exhibits. This process not only helped us identify which design features could be effective in communicating AI concepts but also revealed practical strategies for embedding embodiment and creativity into AI learning experiences more broadly.

In response to RQ2, we conducted a study of visitor learning and interest development, assessed whether our conjectures were supported by our results, and found that different forms of embodiment could support different ways of reasoning about AI concepts across the three exhibits (embodiment takeaway). We also found that creative activities afford connections between AI and self, which in turn could contribute to learners’ interest in AI and their understanding of AI concepts (creativity takeaway). The remainder of

this section will synthesize our findings across our three exhibits, report design recommendations for AI museum exhibits based on our findings, and put our findings in conversation with other work in the field.

6.1 Embodiment Takeaway: Different Forms of Embodiment Can Support Different Ways of Reasoning about AI Concepts.

Across the three exhibits we designed, we incorporated three distinct forms of embodiment—tangible interaction, body knowledge, and full-body interaction (see Table 7). Our data revealed that tangible interaction helped learners grasp AI concepts by making abstract ideas more concrete in the *DataBites* exhibit; however, further refinement is needed to ensure that tangible interactivity is consistently integrated across all phases of the exhibit experience. Full-body interaction, on the other hand, helped learners connect their actions to the AI’s responses in the *LuminAIx* exhibit, though additional research is needed to determine whether this form of interaction effectively supports deeper understanding of the learning objectives. *Knowledge Net* did not sufficiently support body knowledge to assess how it might contribute to learners’ reasoning about AI outputs. Together, these findings suggest that tangibility and full-body interaction may engage different mechanisms for making sense of AI. Future research at the intersection of tangible and embodied interaction and AI education should further investigate how different types of embodiment can influence learners’ understanding and reasoning about AI concepts. Below, we elaborate further on the findings related to tangibility and full-body interaction and their implications for designing AI learning experiences.

6.1.1 Tangibility Makes AI Concepts Approachable and Concrete. Many AI concepts are abstract, especially for young learners. For example, datasets for ML models are typically vast and numerical, making it hard for non-technical learners to identify patterns and explore how data inputs affect AI outputs. We found that utilizing tangibles in *DataBites* aided learners in developing a more concrete understanding of these AI concepts. The physical pizza and sandwich pieces provided an approachable way for participants to manipulate data, as evidenced by them primarily noticing the dataset preparation step. Moreover, instead of providing a static database of pizza and sandwich images, learners can construct and interact with the individual data points, investigating how changing specific features affects the output in a way that is personally relevant. For example, a child in a recorded session created a sandwich by physically stacking ingredients as they would when preparing to eat it. To the top-down camera view, the sandwich just looked like a single piece of bread. The act of assembling the sandwich embodied the child’s understanding of what a sandwich should look like as a data point to teach the computer and allowed the child to concretely explore how the AI’s concept of a sandwich differs from their own. Such tangible interactions allow for exploration and creativity, fostering a richer engagement with the concept of data and its role in AI training. This follows the findings of Zhou et al. [69] that when interacting with a light-based tangible user interface that teaches young learners about AI recommendations, the learners found the learning material more approachable due

Embodiment Implications	Type of Embodiment	Instantiation in Exhibits	Evidence
Tangibility can make AI concepts approachable and concrete , but it needs to be integrated into all the phases of the exhibit experience.	Tangible interaction	<i>DataBites</i> (creation of a training dataset using physical pizza/sandwich pieces)	Most participants (70.83%) identified that dataset preparation, the portion of the exhibit that featured tangible interaction, is part of the ML training process.
We did not find that learners utilized body knowledge in the exhibit , so further work is needed to more closely tie body knowledge to the learning objectives and test its role in learners' understanding.	Using body knowledge	<i>Knowledge Net</i> (humanoid character bodies and motions)	Insufficient evidence
Full-body interaction links learners' actions to AI responses, highlighting AI's human-like and technical aspects , but more work is needed to test how full-body interaction mediates AI understanding.	Full-body interaction	<i>LuminAlx</i> (using dance moves as an input and output medium for interacting with an AI system)	Most participants (66.67%) identified that the AI dance partner responded to their dance moves. Participants who used both mechanization and anthropomorphism when talking about the AI agent scored the highest on average across all questions (n=8, M=2.03).

Table 7: Summary of embodiment types and their implications in each exhibit

to their familiarity with and daily experiences observing light. We suggest that making abstract concepts, such as data, tangible in museum exhibits can foster understanding of these concepts in young learners.

6.1.2 Full-Body Interaction Links Learners' Actions to AI Responses, Highlighting AI's Human-like and Technical Aspects. Full-body interaction enabled learners to connect their own movements to the AI's responses, helping them form deeper conceptualizations of how AI functions as both human-like and technical, ultimately leading to a better understanding of how AI learns from humans. We found that learners who conceptualized AI as having both human (*anthropomorphism*) and non-human (*mechanization*) traits had the highest understanding of our learning objectives for *LuminAlx*. Prior work has found that in augmented reality environments, learners create a *liminal world* in which they blend their physical world and the corresponding scientific world [23]. In our case, learners seem to blend their experience of collaborating with a human-like AI dance partner (their physical world) and the technical components that underlie the AI dance partner (the scientific world). We theorize that exhibit designs that support the creation of liminal worlds can support understanding of AI as both a human-like agent and a technical artifact.

6.2 Creativity Takeaway: Creative Activities Afford Connections Between AI & Self, Contributing to Interest in AI & Understanding of AI.

In our three exhibits, we utilized two forms of creativity: free-choice/personally relevant creation and personalization (see Table 8). Our data demonstrated that in *DataBites* and *LuminAlx*, giving learners freedom to explore different inputs, including personally meaningful ones, to the AI systems engaged them in learning about AI. In *Knowledge Net*, the opportunity for personalization of AI characters engaged learners, which in turn led to better understanding of the AI concepts. Together, these findings demonstrate that creativity in AI learning experiences can foster engagement and situational interest in learning about AI. The following section will provide a more in-depth explanation of the role of personally relevant creative experiences in learning about AI.

6.2.1 Freedom to Create Familiar, Personally Relevant AI Experiences Promotes Interest in AI. In all three exhibits, learners were given the freedom to shape their experiences around different topics. However, participants in *DataBites* and *LuminAlx* had greater control and were able to engage with content that was personally meaningful to them—food and dance, respectively—compared to the *Knowledge Net* exhibit, where customization options were limited and the content was not necessarily connected to familiar, meaningful topics (but still engaged participants in

Creativity Implications	Type of Creativity	Instantiation in Exhibits	Evidence
Freedom to create familiar, personally relevant AI experiences promotes situational interest in AI , such as creating personally relevant artifacts (pizzas and sandwiches) or movements (dance moves).	Free-choice exploration and personally relevant creation	<i>DataBites</i> (learners can create a dataset of pizzas and sandwiches they like), <i>LuminAIx</i> (learners can interact with the AI dance partner by performing dance moves they enjoy)	In <i>DataBites</i> , of the 12 participants who mentioned personally meaningful creation, over half (n=7, 58.3%) expressed that they wanted to learn more about AI in the future. In <i>LuminAIx</i> , after dancing with the AI agent, a large majority of participants expressed interest in learning more about AI (n=14, 77.77%).
Opportunity to personalize AI learning can foster engagement, contributing to AI understanding , but we should update the design to further encourage learners to try different inputs.	Personalization	<i>Knowledge Net</i> (learners can customize their characters by modifying a semantic network)	Of the participants who engaged in personalization by trying a variety of inputs to their networks, all of them (n=9, 100%) demonstrated partial understanding of the learning objectives.

Table 8: Summary of creativity types and their implications in each exhibit

the learning activity). In *DataBites*, participants could personalize food items, experimenting with different combinations as part of their learning process, while in *LuminAIx*, they could explore dance movements in ways that felt expressive and unique. Qualitative coding of participants' responses, along with their self-reported interest in AI, indicated high levels of engagement and interest development in both *DataBites* and *LuminAIx*. Overall, these findings suggest that AI learning experiences that allow learners to interact with familiar, meaningful topics and shape experiences that are personally relevant can effectively foster sustained interest in learning about AI.

These findings corroborate Belghith et al.'s assertion that AI learning experiences can benefit from incorporating learners' interests [13] and Dangol et al.'s recommendation that AI learning experiences should engage learners' cultural knowledge [17]. Moreover, this aligns with the findings of Kim et al. [35] that playful experiences that center learners' lived experiences, interests, and culture can make learning about AI more accessible and enjoyable. Our findings motivate future AI museum exhibits—and AI education initiatives more broadly— which center on open-ended, creative activities where learners can incorporate personally relevant ideas.

7 Limitations and Future Work

This study has several limitations which suggest directions for future work. Our sample size and single museum study restrict the generalizability of our findings and may not capture the full range of visitor interactions and perspectives. We chose to evaluate the exhibits in-situ, which prevented us from studying the impact of individual variables as one would do in a controlled study. The exhibit prototypes themselves represent early design iterations with some usability issues. Further development is planned to address these challenges, corroborate our findings with additional studies,

and enhance the overall visitor experience based on insights from this study.

Our evaluation instruments were also in their initial phase of development. While they provided valuable insights into participant learning and interest development, we identified areas for improvement (e.g., ensuring that the content knowledge questions do not evaluate knowledge of details outside the scope of the learning objectives). We also encountered issues with assessing creativity and embodiment. Since the data analyzed for these factors came from responses not specifically designed to evaluate creativity and embodiment, the findings highlighted salient features but lacked consistency across exhibits. Moving forward, we plan to develop targeted instruments to evaluate these constructs. These tools will be standardized to contribute to the TEI and AI education research communities.

Finally, to evaluate future iterations of the exhibits, we will conduct an in-depth video analysis examining how creativity, embodiment, and learning talk [56] occur in visitor interactions. This analysis will deepen our understanding of how embodiment and creativity foster discussion about the learning content of our exhibits.

8 Conclusion

The primary contribution of this paper is the design of three exhibit prototypes—*DataBites*, *Knowledge Net*, and *LuminAIx*—that aim to promote AI literacy and foster middle schoolers' interest in AI through embodiment and creativity. We present findings from interviews with 29 10-14 year old visitors, assessing their conceptual understanding, interest development, and how salient embodiment and creativity were to their experiences. We also present a multi-case analysis of how embodiment and creativity were leveraged to support AI learning across three exhibits. Our results suggest that AI museum exhibits that leverage free-choice and personally

meaningful activities can engage learners and support their interest in AI; additionally, we found that different forms of embodiment can contribute to different ways of reasoning about AI. However, our exhibits need future refinement to communicate their learning objectives and to address challenges related to usability and evaluation in museum settings. Additionally, we recommend that AI learning experiences in museums leverage creative and personally relevant activities, support hybrid conceptualizations of AI agents, and utilize tangible interfaces to make AI concepts approachable and concrete. This paper offers insights into designing interfaces that support young learners' AI literacy, and it also provides valuable guidance for educators and museum exhibit designers.

Acknowledgments

We would like to thank our exhibit designers and developers with whom this research would not be possible—Jasmin Ali-Diaz, Dev Ambani, Chelsi Cocking, Katherine Hancock, Ramya Iyer, Cassandra Naomi, Nyssa Shahdadhuri, Gemma Tate, and Harisri Yeramsetti. We would also like to thank Molly de Blanc, Emily Dobar, and Atefeh Mahdavi Goloujeh for their support during our museum study. Thank you to the staff of Griffin Museum of Science and Industry, especially Dr. Lauren Applebaum, for allowing us to test our exhibits in their space. Thank you to our project advisory board—Dr. Michael Horn (who we also thank for lending us his Ideum table), Rob Lindgren, and Fred Martin. Finally, thank you to our anonymous reviewers who provided valuable feedback that helped us strengthen our paper. This work was funded by the National Science Foundation (DRL 2214463).

References

- [1] Sasha A. Barab (Ed.). 2016. *Design-based Research: Clarifying the Terms. A Special Issue of the Journal of the Learning Sciences* (Mahwah, 2016). Number volume 13, number 1 in The journal of the learning sciences. Taylor and Francis Group.
- [2] 2025. *Exploring AI: Making the Invisible Visible | Museum of Science*. <https://www.mos.org/visit/exhibits/exploring-ai-making-invisible-visible>
- [3] Thomas Humphrey, Joshua P. Gutwill, and Exploratorium (Organization) (Eds.). 2005. *Fostering active prolonged engagement: the art of creating APE exhibits* (San Francisco, 2005). Exploratorium. OCLC: ocm61162170.
- [4] 2023. *Virtually Human Explores What it Means for AI to Act Like a Human*. <https://lawrencehallofscience.org/news/virtually-human-explores-what-it-means-for-ai-to-act-like-a-human/>
- [5] Dor Abrahamson and Arthur Bakker. 2016. Making sense of movement in embodied design for mathematics learning. 1, 1 (2016), 33. doi:10.1186/s41235-016-0034-3
- [6] Ibrahim Oluwajoba Adisa, Ian Thompson, Tolulope Famaye, Deepika Sistla, Cinnamon Sunrise Bailey, Katherine Mulholland, Alison Fecher, Caitlin Marie Lancaster, and Golnaz Arastoopour Irgens. 2023. S.P.O.T: A Game-Based Application for Fostering Critical Machine Learning Literacy Among Children. In *Proceedings of the 22nd Annual ACM Interaction Design and Children Conference* (Chicago IL USA, 2023-06-19). ACM, 507–511. doi:10.1145/3585088.3593884
- [7] Sue Allen. 2004. Designs for learning: Studying science museum exhibits that do more than entertain. 88 (2004), S17–S33. Issue S1. doi:10.1002/sce.20016
- [8] Alissa N. Antle, Greg Corness, and Allen Bevans. 2013. Balancing justice: comparing whole body and controller-based interaction for an abstract domain. 6, 4 (2013), 388. doi:10.1504/IJART.2013.058285
- [9] Alissa N. Antle, Greg Corness, and Milena Droumeva. 2009. Human-computer-intuition? Exploring the cognitive basis for intuition in embodied interaction. 2, 3 (2009), 235. doi:10.1504/IJART.2009.028927
- [10] Patricia Armstrong. [n. d.]. *Bloom's Taxonomy*. <https://cft.vanderbilt.edu/guides-sub-pages/blooms-taxonomy/>
- [11] Saskia Bakker, Alissa N. Antle, and Elise Van Den Hoven. 2012. Embodied metaphors in tangible interaction design. 16, 4 (2012), 433–449. doi:10.1007/s00779-011-0410-4
- [12] Leslie Bedford. 2014. *The art of museum exhibitions: how story and imagination create aesthetic experiences*. Left Coast Press.
- [13] Yasmine Belghith, Atefeh Mahdavi Goloujeh, Brian Magerko, Duri Long, Tom Mcklin, and Jessica Roberts. 2024. Testing, Socializing, Exploring: Characterizing Middle Schoolers' Approaches to and Conceptions of ChatGPT. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (Honolulu HI USA, 2024-05-11). ACM, 1–17. doi:10.1145/3613904.3642332
- [14] Margaret A Boden. 2004. *The creative mind: Myths and mechanisms*. Routledge.
- [15] Leah Buechley, Mike Eisenberg, Jaime Catchen, and Ali Crockett. 2008. The LilyPad Arduino: using computational textiles to investigate engagement, aesthetics, and diversity in computer science education. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Florence Italy, 2008-04-06). ACM, 423–432. doi:10.1145/1357054.1357123
- [16] Francesco Cafaro and Jessica Roberts. 2021. *Data through Movement: Designing Embodied Human-Data Interaction for Informal Learning* (1st ed. 2021 ed.). Springer International Publishing.
- [17] Aayushi Dangol, Michele Newman, Robert Wolfe, Jin Ha Lee, Julie A. Kientz, Jason Yip, and Caroline Pitt. 2024. Mediating Culture: Cultivating Socio-cultural Understanding of AI in Children through Participatory Design. In *Designing Interactive Systems Conference* (IT University of Copenhagen Denmark, 2024-07). ACM, 1805–1822. doi:10.1145/3643834.3661515
- [18] Hasti Darabipourshiraz, Dev Ambani, and Duri Long. 2024. DataBites: An embodied and co-creative museum exhibit to foster children's understanding of supervised machine learning. In *Proceedings of the 16th Conference on Creativity & Cognition* (Chicago, IL, USA) (C&C '24). Association for Computing Machinery, New York, NY, USA, 550–555. doi:10.1145/3635636.3664247
- [19] Paul Dourish. 2004. *Where the Action Is: The Foundations of Embodied Interaction*. MIT Press.
- [20] Stefania Druga. 2023. Creative AI Literacies for Families. <https://digital.lib.washington.edu/researchworks/items/03853bd4-b257-4fe8-8158-b08032c5efd8/full>
- [21] Stefania Druga, Fee Lia Christoph, and Amy J Ko. 2022. Family as a Third Space for AI Literacies: How do children and parents learn about AI together?. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New York, NY, USA, 2022-04-29) (CHI '22). Association for Computing Machinery, 1–17. doi:10.1145/3491102.3502031
- [22] Stefania Druga, Sarah T. Vu, Eesh Likhith, and Tammy Qiu. 2019. Inclusive AI literacy for kids around the world. In *Proceedings of FabLearn 2019* (New York NY USA, 2019-03-09). ACM, 104–111. doi:10.1145/3311890.3311904
- [23] Noel Enyedy, Joshua A. Danish, and David DeLiema. 2015. Constructing liminal blends in a collaborative augmented-reality learning environment. 10, 1 (2015), 7–34. doi:10.1007/s11412-015-9207-1
- [24] John H. Falk and Lynn D. Dierking. 2018. *Learning from museums* (second edition ed.). Rowman & Littlefield.
- [25] Alan J. Friedman, Sue Allen, Patricia B. Campbell, Lynn D. Dierking, Barbara N. Flagg, Cecilia Garibay, Randi Korn, Gary Silverstein, and David A. Ucko. 2008. Framework for Evaluating Impacts of Informal Science Education Projects. http://insci.org/resources/Eval_Framework.pdf
- [26] Vanessa Graf. 2019. *Understanding AI – a place to experience artificial intelligence – Ars Electronica Blog*. <https://ars.electronica.art/aeblog/en/2019/08/06/understanding-ai-futurelab-installations/>
- [27] Eric Greenwald, Ari Krakowski, Timothy Hurt, Kelly Grindstaff, and Ning Wang. 2024. It's like I'm the AI: Youth Sensemaking About AI through Metacognitive Embodiment. In *Proceedings of the 23rd Annual ACM Interaction Design and Children Conference* (Delft Netherlands, 2024-06-17). ACM, 789–793. doi:10.1145/3628516.3659395
- [28] Suzanne Hidi and K. Ann Renninger. 2006. The Four-Phase Model of Interest Development. 41, 2 (2006), 111–127. doi:10.1207/s15326985ep4102_4
- [29] Michael Horn, Zeina Atrash Leong, Florian Block, Judy Diamond, E. Margaret Evans, Brenda Phillips, and Chia Shen. 2012. Of BATs and APes: an interactive tabletop game for natural history museums. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, Austin Texas USA, 2059–2068. doi:10.1145/2207676.2208355
- [30] Eva Hornecker and Jacob Buur. 2006. Getting a grip on tangible interaction: a framework on physical space and social interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Montréal Québec Canada, 2006-04-22). ACM, 437–446. doi:10.1145/1124772.1124838
- [31] Eva Hornecker and Luigina Ciolfi. 2019. *Human-Computer Interactions in Museums*. Springer International Publishing. doi:10.1007/978-3-031-02225-8
- [32] Mark Johnson. [n. d.]. *The body in the mind: the bodily basis of meaning, imagination, and reason* (paperback ed., [9. print.] ed.). University of Chicago Press.
- [33] Brian Jordan, Nisha Devasia, Jenna Hong, Randi Williams, and Cynthia Breazeal. 2021. PoseBlocks: A Toolkit for Creating (and Dancing) with AI. *Proceedings of the AAAI Conference on Artificial Intelligence* 35, 17 (May 2021), 15551–15559. doi:10.1609/aaai.v35i17.17831
- [34] Wataru Kawabe, Yuri Nakao, Akihisa Shitara, and Yusuke Sugano. 2024. Technical Understanding from Interactive Machine Learning Experience: a Study Through a Public Event for Science Museum Visitors. 36, 3 (2024), 155–171. doi:10.1093/iwc/iwae007

- [35] Yoon Jeon Kim, Gyeongri Kim, and Andy Stoiber. 2025. Playful design for AI Literacy: Creating inclusive learning and assessment opportunities. *Int. J. Hum.-Comput. Stud.* 199, C (May 2025). doi:10.1016/j.ijhcs.2025.103508
- [36] George Lakoff and Mark Johnson. 2003. *Metaphors we live by*. University of Chicago Press.
- [37] Sunok Lee, Dasom Choi, Minha Lee, Jonghak Choi, and Sangsu Lee. 2023. Fostering Youth's Critical Thinking Competency About AI through Exhibition. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23)*. Association for Computing Machinery, New York, NY, USA, 1–22. doi:10.1145/3544548.3581159
- [38] Scott K. Liddell. 1998. Grounded blends, gestures, and conceptual shifts. 9, 3 (1998), 283–314. doi:10.1515/cogl.1998.9.3.283
- [39] Robb Lindgren and Mina Johnson-Glenberg. 2013. Emboldened by Embodiment: Six Precepts for Research on Embodied Learning and Mixed Reality. 42, 8 (2013), 445–452. doi:10.3102/0013189X13511661
- [40] Duri Long, Takeria Blunt, and Brian Magerko. 2021. Co-Designing AI Literacy Exhibits for Informal Learning Spaces. 5 (2021), 1–35. Issue CSCW2. doi:10.1145/3476034
- [41] Duri Long and Brian Magerko. 2020. What is AI Literacy? Competencies and Design Considerations. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (Honolulu HI USA, 2020-04-21)*. ACM, 1–16. doi:10.1145/3313831.3376727
- [42] Duri Long, Tom McKlin, Anna Weisling, William Martin, Hannah Guthrie, and Brian Magerko. 2019. Trajectories of Physical Engagement and Expression in a Co-Creative Museum Installation. In *Proceedings of the 2019 on Creativity and Cognition (San Diego CA USA, 2019-06-13)*. ACM, 246–257. doi:10.1145/3325480.3325505
- [43] Duri Long, Aadarsh Padiyath, Anthony Teachey, and Brian Magerko. 2021. The Role of Collaboration, Creativity, and Embodiment in AI Learning Experiences. In *Proceedings of the 13th Conference on Creativity and Cognition (C&C '21)*. Association for Computing Machinery, New York, NY, USA, 1–10. doi:10.1145/3450741.3465264
- [44] Duri Long, Sophie Rollins, Jasmin Ali-Diaz, Katherine Hancock, Samnang Nuonsinoeun, Jessica Roberts, and Brian Magerko. 2023. Fostering AI Literacy with Embodiment & Creativity: From Activity Boxes to Museum Exhibits. In *Proceedings of the 22nd Annual ACM Interaction Design and Children Conference (New York, NY, USA, 2023-06-19) (IDC '23)*. Association for Computing Machinery, 727–731. doi:10.1145/3585088.3594495
- [45] Duri Long, Anthony Teachey, and Brian Magerko. 2022. Family Learning Talk in AI Literacy Learning Activities. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems (New York, NY, USA, 2022-04-29) (CHI '22)*. Association for Computing Machinery, 1–20. doi:10.1145/3491102.3502091
- [46] Duri Long, Jiaxi Yang, Cassandra Naomi, and Brian Magerko. 2024. Xylocode: A Novel Approach to Fostering Interest in Computer Science via an Embodied Music Simulation. In *Proceedings of the CHI Conference on Human Factors in Computing Systems (Honolulu HI USA, 2024-05-11)*. ACM, 1–16. doi:10.1145/3613904.3642098
- [47] Leilah Lyons, Brian Slattery, Priscilla Jimenez, Brenda Lopez, and Tom Moher. 2012. Don't forget about the sweat: effortful embodied interaction in support of learning. In *Proceedings of the Sixth International Conference on Tangible, Embedded and Embodied Interaction (Kingston Ontario Canada, 2012-02-19)*. ACM, 77–84. doi:10.1145/2148131.2148149
- [48] Brian Magerko, Jason Freeman, Tom Mcklin, Mike Reilly, Elise Livingston, Scott Mccoid, and Andrea Crews-Brown. 2016. EarSketch: A STEAM-Based Approach for Underrepresented Populations in High School Computer Science Education. 16, 4 (2016), 1–25. doi:10.1145/2886418
- [49] Laura Malinverni and Narcis Pares. 2014. Learning of Abstract Concepts through Full-Body Interaction: A Systematic Review. 17, 4 (2014), 100–116. <https://www.jstor.org/stable/jeductechsoci.17.4.100>
- [50] Paul Marshall. 2007. Do tangible interfaces enhance learning?. In *Proceedings of the 1st international conference on Tangible and embedded interaction (Baton Rouge Louisiana, 2007-02-15)*. ACM, 163–170. doi:10.1145/1226969.1227004
- [51] Luca Moser. 2021. Design of an Easy-to-Use Mobile Augmented Reality Learning Setting by Means of a Conjecture Map. *International Association for Development of the Information Society* (2021).
- [52] Seymour Papert. 1993. *Mindstorms: children, computers, and powerful ideas* (2nd ed.). Basic Books.
- [53] Kim Patch. 2022. *MIT's FutureMakers programs help kids get their minds around — and hands on — AI*. <https://medium.com/mit-media-lab/mits-futuremakers-programs-help-kids-get-their-minds-around-and-hands-on-ai-c426737a9c7>
- [54] William Christopher Payne, Yoav Bergner, Mary Etta West, Carlie Sharp, R. Benjamin Shapiro, Danielle Albers Szafir, Edd V. Taylor, and Kayla DesPortes. 2021. danceON: Culturally Responsive Creative Computing. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (Yokohama Japan, 2021-05-06)*. ACM, 1–16. doi:10.1145/3411764.3445149
- [55] Janet C. Read and Stuart MacFarlane. 2006. Using the fun toolkit and other survey methods to gather opinions in child computer interaction. In *Proceedings of the 2006 conference on Interaction design and children (Tampere Finland, 2006-06-07)*. ACM, 81–88. doi:10.1145/1139073.1139096
- [56] Jessica Roberts and Leilah Lyons. 2017. The value of learning talk: applying a novel dialogue scoring method to inform interaction design in an open-ended, embodied museum exhibit. 12, 4 (2017), 343–376. doi:10.1007/s11412-017-9262-x
- [57] Sophie Rollins, Katherine Hancock, Jasmin Ali-Diaz, Nyssa Shahdarpuri, and Duri Long. 2024. Knowledge Net: Fostering Children's Understanding of Knowledge Representations Through Creative Making and Embodied Interaction in a Museum Exhibit. In *Proceedings of the 16th Conference on Creativity & Cognition (Chicago, IL, USA) (C&C '24)*. Association for Computing Machinery, New York, NY, USA, 470–475. doi:10.1145/3635636.3664254
- [58] William Sandoval. 2014. Conjecture Mapping: An Approach to Systematic Educational Design Research. 23, 1 (2014), 18–36. doi:10.1080/10508406.2013.778204
- [59] William A. Sandoval and Philip Bell. 2004. Design-Based Research Methods for Studying Learning in Context: Introduction. 39, 4 (2004), 199–201. doi:10.1207/s15326985ep3904_1
- [60] Scott S. Snibbe and Hayes S. Raffle. 2009. Social immersive media: pursuing best practices for multi-user interactive camera/projector exhibits. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (Boston MA USA, 2009-04-04)*. ACM, 1447–1456. doi:10.1145/1518701.1518920
- [61] Elisabeth Sulmont, Elizabeth Patitsas, and Jeremy R. Cooperstock. 2019. Can You Teach Me To Machine Learn?. In *Proceedings of the 50th ACM Technical Symposium on Computer Science Education (Minneapolis MN USA, 2019-02-22)*. ACM, 948–954. doi:10.1145/3287324.3287392
- [62] David Touretzky, Christina Gardner-McCune, Fred Martin, and Deborah Seehorn. 2019. Envisioning AI for K-12: What Should Every Child Know about AI? 33, 1 (2019), 9795–9799. doi:10.1609/aaai.v33i01.33019795
- [63] Steven Van Wolputte. 2004. Hang on to Your Self: Of Bodies, Embodiment, and Selves. 33 (2004), 251–269. <https://www.jstor.org/stable/25064853>
- [64] Iro Voulgari, Marvin Zammit, Elias Stouraitis, Antonios Liapis, and Georgios Yannakakis. 2021. Learn to Machine Learn: Designing a Game Based Approach for Teaching Machine Learning to Primary and Secondary Education Students. In *Interaction Design and Children (Athens Greece, 2021-06-24)*. ACM, 593–598. doi:10.1145/3459990.3465176
- [65] Randi Williams. 2023. A review of assessments in k-12 ai literacy curricula. Retrieved October 25 (2023), 2023.
- [66] Randi Williams, Safinah Ali, Raúl Alcantara, Tasneem Burghleh, Sharifa Alghowinem, and Cynthia Breazeal. 2024. Doodlebot: An Educational Robot for Creativity and AI Literacy. In *Proceedings of the 2024 ACM/IEEE International Conference on Human-Robot Interaction (Boulder CO USA, 2024-03-11)*. ACM, 772–780. doi:10.1145/3610977.3634950
- [67] Chengzhi Zhang, Chelsi Alise Cocking, Milka Trajkova, Zoe Lacy Mock, Gemma Tate, Cassandra Naomi Monden, and Brian Magerko. 2024. Fostering AI Literacy with LuminAI through Embodiment and Creativity in Informal Learning Spaces. In *Proceedings of the 16th Conference on Creativity & Cognition (Chicago, IL, USA) (C&C '24)*. Association for Computing Machinery, New York, NY, USA, 476–481. doi:10.1145/3635636.3664255
- [68] Helen Zhang, Irene Lee, Safinah Ali, Daniella DiPaola, Yihong Cheng, and Cynthia Breazeal. 2023. Integrating Ethics and Career Futures with Technical Learning to Promote AI Literacy for Middle School Students: An Exploratory Study. 33, 2 (2023), 290–324. doi:10.1007/s40593-022-00293-3
- [69] Xiaofei Zhou, Yi Zhang, Yufei Jiang, Yunfan Gong, Chi Zhang, Alissa N. Antle, and Zhen Bai. 2025. Briteller: Shining a Light on AI Recommendations for Children. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (CHI '25)*. Association for Computing Machinery, New York, NY, USA, 1–30. doi:10.1145/3706598.3714106
- [70] Xiaofei Zhou, Yushan Zhou, Yunfan Gong, Zhenyao Cai, Annie Qiu, Qinqin Xiao, Alissa N. Antle, and Zhen Bai. 2024. "Bee and I need diversity!" Break Filter Bubbles in Recommendation Systems through Embodied AI Learning. In *Proceedings of the 23rd Annual ACM Interaction Design and Children Conference (New York, NY, USA, 2024-06-17) (IDC '24)*. Association for Computing Machinery, 44–61. doi:10.1145/3628516.3655802
- [71] Tom Ziemke. 2003. What's that Thing Called Embodiment? 25, 25 (2003). <https://escholarship.org/uc/item/60w6v9jz>
- [72] Heather Toomey Zimmerman, Susan M. Land, Katharine Grills, Yu-Chen Chiu, Yong Ju Jung, and Jeffrey Williams. 2020. Design Conjectures for Place-Based Science Learning about Water: Implementing Mobile Augmented Reality with Families. (June 2020). <https://repository.isls.org/handle/1/6304> Publisher: International Society of the Learning Sciences (ISLS).