Interdisciplinary Practice in Industrial Design, Vol. 222, 2026, 305–315
https://doi.org/10.54941/ahfe1007769

Music Track Story: Children's Music Track Rolling Ball Building Blocks Based on Physical Acoustics and AIoT

Chanwen Pu

Zhejiang Sci-Tech University, Hangzhou, ZJ 311103, China

ABSTRACT

This paper investigates the application of AIoT (Artificial Intelligence of Things) in designing musical early-learning toys for children within the STEAM education framework. By integrating principles of physical acoustics with intelligent interaction, the research aims to create novel and intuitive play experiences. Grounded in a theoretical review of STEAM and acoustics, this paper employs case analysis to evaluate existing products across three key dimensions related to human factors: technological integration, interactive diversity, and user engagement. The analysis identifies specific AIoT application models suitable for children's products. In the design phase, the study proposes a modular system that incorporates AI-assisted creation tools and IoT connectivity. This system culminates in the development of interactive music blocks and track pieces that allow for the free combination of scales and tones. Through tangible manipulation and real-time auditory feedback, children can explore musical structures in a playful, low-pressure environment. The proposed framework offers new insights into how smart educational toys can facilitate implicit learning, contributing to the integrated development of toy design and educational technology.

Keywords: STEAM education, Toy design, Music enlightenment, AIoT

INTRODUCTION

Under the STEAM education philosophy, early childhood education advocates for natural learning through authentic experiences. As a classic toy, building blocks offer substantial educational potential during spatial construction. For instance, ball-track blocks incorporate physics concepts such as gravity and motion, yet they rarely translate these physical interactions into explicit learning opportunities. Similarly, while musical toys can produce melodies, they often remain confined to the level of "sound generation," failing to help children grasp the underlying mechanisms of core musical concepts like pitch, timbre, and rhythm. Although each toy category touches upon specific facets of STEAM, they fall short of truly integrating "physical motion" with "sound production" to achieve synergistic cross-disciplinary learning.

Existing research indicates that building block toys promote spatial thinking development (Brosterman, 1997). As a classic example, LEGO construction activities not only enhance children's building abilities but also produce near-transfer effects on arithmetic skills, suggesting that explicitly

Received March 6, 2026; Revised April 5, 2026; Accepted April 20, 2026; Available online July 20, 2026

integrating spatial strategies into construction tasks holds potential value for mathematical learning (McDougal et al., 2024). Concurrently, the positive impact of music games on auditory perception has been empirically supported. Gamified learning in VR rhythm games significantly enhances children's rhythmic perception and performance abilities (Pesek et al., 2024). Furthermore, research on music education systems incorporating motion-based interaction and game elements indicates that gamified music learning effectively stimulates children's interest in music and enriches their auditory experiences (Sun, 2025).

However, research on these two domains rarely intersects. The few products attempting to combine them often merely overlay electronic sound effects onto physical structures, lacking an intrinsic logic linking music and physical space. As a result, their educational value remains superficial. How to make physical spatial structures influence musical attributes and how to enable children to naturally encounter music theory knowledge while constructing tracks remain significant challenges in current children's product design.

Today, the rapid advancement of AIoT (Artificial Intelligence of Things) technology has expanded its educational applications, enabling the intelligent upgrading of traditional toys. For example, an AIoT educational technology solution centered on the Firmata protocol and MediaPipe, which was developed through four years of teaching practice, has effectively lowered the technical barrier for learners to achieve AI-driven IoT control (Takefuji, 2025). Nevertheless, systematic design research is still lacking on how these technologies can be organically integrated into children's musical educational toys and combined with physical acoustic principles to realize the educational goal of "learning through play."

Therefore, this study aims to explore how AIoT technology can empower the design of children's musical educational toys. By leveraging physical acoustic principles and intelligent interaction modes, it seeks to provide children with a novel play experience within the STEAM education framework. To achieve this objective, this paper primarily employs case analysis. Through a systematic review of existing physically triggered musical toys, smart-guided early learning products, and interconnected collaborative children's devices, it distills core design elements for AIoT-empowered children's musical educational toys. Building on this foundation, it proposes a design framework for "Music Track Rolling Ball Building Blocks."

RESEARCH BACKGROUND

Building Blocks and Spatial Thinking Development

As a classic construction toy, building blocks have long been recognized for their educational value. The "Gifts" system developed by Froebel, regarded as the prototype of modern building blocks, inherently embodies the idea that abstract concepts can be understood through physical manipulation (Brosterman, 1997). Research has shown that block construction activities promote the development of children's spatial representation abilities, including core skills such as spatial visualization, mental rotation, and the understanding of spatial relationships, skills that lay the foundation for later learning in mathematics and science. Randomized controlled trials have found that free block play significantly enhances children's geometric abilities and behavioral self-regulation, highlighting the critical role of early spatial experiences in supporting STEAM learning (Schmitt et al., 2025). Moreover, related research reveals a significant positive correlation between the complexity of block structures built by children aged 3 to 5 years and their spatial skills one year later (Lauer et al., 2023).

Music Enlightenment and Children's Cognitive Development

The value of music education for early childhood development is well documented. Music-based activities enhance children's auditory discrimination, rhythmic perception, and pitch recognition, skills that are positively correlated with language development and mathematical thinking. A five-year longitudinal study of children aged 6 to 7 years found that children receiving music training outperformed controls in both auditory discrimination and attention control tasks (Habibi et al., 2022). Experimental research further indicates that rhythm training significantly improves preschoolers' performance in language processing tasks, particularly in grammatical awareness (Frischen et al., 2022). By enhancing the precision of the auditory nervous system, music training exerts measurable effects on children's speech processing, reading abilities, and learning capacity. These effects extend to the development of executive functions and verbal memory (Miendlarzewska & Trost, 2014). At the neurobiological level, early musical engagement activates the auditory cortex and promotes neural connectivity, processes associated with attention, memory, and creativity. Rhythm training, in particular, supports physical coordination and self-expression through the synchronization of body movements with musical beats. Moreover, the emotional expression and communication inherent in musical activities are linked to children's empathy development and emotional regulation skills.

Application of AIoT Technology in Educational Toys

The convergence of artificial intelligence (AI) and the Internet of Things (IoT) has opened new avenues for the intelligent upgrading of educational toys. In recent years, smart toys equipped with perception, interaction, and learning capabilities have emerged through the integration of AI into children's playthings. Leveraging natural language processing, computer vision, and adaptive algorithms, these products enable dynamic and personalized interactions with children (Xiao & Gonçalves, 2025). At the IoT level, features such as device interconnectivity, data synchronization, and cloud collaboration allow toys to transcend the limitations of single-user operation, facilitating multiplayer collaborative play scenarios (Rajkumar et al., 2025). The core value of AIoT technology lies in its capacity to transform traditional physical toys into intelligent terminals with perception, feedback, and connectivity capabilities, thereby expanding the educational dimension of toys while preserving the physical play experience. However, how to organically integrate these technologies into the design of children's musical early learning toys, and how to combine them with physical acoustic principles to achieve the educational goal of "learning through play," remain key areas requiring further exploration in both product design and educational research.

RESEARCH METHODOLOGY

This study employs a case analysis approach to explore how AIoT technology can empower the design of children's musical early learning toys. Grounded in the principles of STEAM education and physical acoustics, the study establishes an analytical framework structured around three key dimensions: technological maturity, interactive diversity, and depth of user engagement. Three representative existing products and research cases are selected as the analytical samples, and their design characteristics are examined in depth, focusing on three aspects: physical-musical mapping, intelligent guidance mechanisms, and modes of social collaboration.

Physical-Triggered Cases: Musical Exploration in Ball-Track Toys

As a classic form of constructional building blocks, ball-track toys have been extensively studied for their educational value. In recent years, researchers have begun exploring the potential to translate physical motion into musical feedback. The movement of marbles along tracks inherently exhibits rich acoustic characteristics. Previous studies have achieved real-time sound feedback based on physical models by simulating the sounds generated during marble motion, thereby providing a theoretical foundation for understanding the mapping relationship between physical movement and sound generation (Rath, 2003). With their inherent kinetic properties, ball-track toys serve as a key medium for this exploration. Through the free assembly of track modules with varying heights, curvatures, and lengths, the movement of balls incorporates diverse physical variables, thus establishing a foundation for linking physical manipulation to musical generation.

Taking the HABA Kullerbu Ball Runs as an example, this product utilizes wooden and plastic tracks. Through the free assembly of track modules with varying heights, curvatures, and lengths, combined with sound-producing accessories such as xylophones and bells, the rolling ball triggers pre-set tones via physical collision when passing over xylophone or bell bridges. From the perspective of physical-musical mapping, this design establishes an initial connection between physical actions and sound output. Users can arrange the sequence and placement of sound-producing components; however, the overall design incorporates limited mapping elements for musical expression, resulting in restricted control over parameters such as timbre and rhythm. While physically triggered toys offer a more intuitive perception of sound generation through physical collision, existing products still face significant limitations in establishing mappings between physical variables, such as size, material, and structure, and musical attributes like pitch, timbre, and rhythm.

Figure 1: Germany's HABA kullerbu ball runs.

Smart-Guided Cases: Sound Interaction in AI Emotional Companion Toys

The core value of smart guidance products lies in providing personalized and adaptive learning support. When intelligent systems interact with children as companions rather than instructors, children's learning motivation significantly increases (Belpaeme et al., 2018).

A representative example of this approach is the TalenPal AI Player, developed by Besitech for the global market. This screen-free product comprises an audio base station and multiple interchangeable character dolls known as Talens. Each doll contains themed audio content, such as bedtime stories, nature exploration, or emotional intelligence development. Children can simply place a doll on the base to play its corresponding audio. An AI chat button located on top of the device allows children to engage in conversation with the dolls. A blue voice button is used for exploring sound effects, and a yellow knob enables switching between stories. The system supports real-time AI-powered audio updates, with most content playable offline. Additionally, parents can co-create stories through a mobile application and synchronize them with the player. The design of TalenPal embodies the core concept of intelligent companionship, wherein AI expands children's exploratory boundaries through interactive dialogue. However, the product's interaction primarily relies on voice and auditory modalities, positioning children primarily as listeners and conversational participants. Their auditory experience largely depends on AI-generated output, lacking the physical, hands-on engagement in which sound is actively produced through manual actions. Therefore, a key direction for future exploration lies in how to retain the benefits of AI companionship while enabling children to actively generate and shape sound or even music through physical manipulation. This represents a promising avenue for integrating smart guidance technology with physical construction play.

Figure 2: TalenPal AI player product concept diagram.

Interconnected Collaborative Cases: The LEGO Smart Play System

The LEGO Smart Play System represents a cutting-edge exploration in connected and collaborative toys. This system utilizes a building block kit featuring embedded SMART bricks, which integrate custom ASIC chips and positioning systems. These bricks can precisely detect the relative positions of nearby SMART bricks, tags, and minifigures; they react to each other's presence and recognize whether they are being twisted, swung, or thrown. For auditory feedback, the system employs synthetic soundscapes technology, which decomposes core sounds into fundamental frequencies and amplitudes. By adjusting parameters, it generates entirely distinct sound effects, enabling bricks to mimic vehicles, spacecraft, animals, characters, and more, with virtually unlimited sound possibilities. Crucially, the system supports the interconnection of multiple brick sets, allowing children's creations to recognize and respond to one another, thereby achieving true collaborative play. This systematic design addresses three core needs of children: social interaction with peers, responsive engagement through action, and evolving play experiences that adapt to gameplay. True creative learning occurs during collaborative construction, where children co-design, debug, and share creations with peers, resulting in a depth of learning that far surpasses individual exploration (Resnick, 2017).

This case study reveals that when toys possess interconnected collaborative capabilities, children's exploration extends from individual construction to social co-creation. In this context, sound transforms from a solitary activity into a collaborative ensemble composed with peers. At the same time, technology is cleverly concealed within the building blocks, ensuring that children remain focused on the physical construction experience. However, existing interconnected collaborative toys still exhibit a disconnect between their social features and learning objectives. How to align social interaction with physical and musical learning goals remains an unresolved challenge in design practice.

Figure 3: LEGO "SMART Play" modules.

Case Comparison and Design Implications

A systematic review of the three case studies reveals that establishing an intuitive mapping between physical movement and music generation is essential for achieving embodied learning. However, existing products tend to rely on relatively limited mapping dimensions. Although intelligent guidance and companionship preserve children's immersive play experience more effectively than directive instruction, current technologies still require meaningful integration with physical gameplay. When toys incorporate social attributes, children's exploration shifts from individual to collaborative modes, significantly amplifying the learning value. Nevertheless, the synergies between social interaction and learning objectives require further refinement. The "Music Rolling Ball Track Building Blocks" design framework proposed in this study aims to address the limitations identified in the aforementioned cases through exploratory designs in three key areas: the physical-to-musical mapping mechanism, the integration approach for intelligent guidance, and the implementation pathways for social collaboration.

PRODUCT SYSTEM DESIGN: MUSIC TRACK ROLLING BALL BUILDING BLOCKS

Based on the core design elements extracted from the case studies, this study proposes a design framework for Musical Rolling Ball Track Building Blocks. The system comprises two main modules and is designed to achieve an intrinsic integration between physical construction and music generation.

Physical-Music Mapping Module

The Physical-Musical Mapping Module serves as the foundational component of this design. Its conceptual basis is informed by physically triggered case studies, which demonstrate that establishing an intuitive mapping between physical motion and musical generation is essential for achieving embodied learning.

The building block set includes balls of varying material densities and interchangeable collision surfaces, such as wood and metal, which are paired with percussion modules corresponding to different musical instruments to shape diverse timbres. Children can perceive tonal differences by swapping collision surfaces and instrument blocks. In addition, variations in block structure, such as hollow versus solid forms, and differences in collision shape further influence timbre, offering rich opportunities for exploration. Consider the tone brick module within the instrument blocks: bricks of different lengths correspond to distinct pitches. To achieve desired pitch effects, children naturally experiment with selecting tone bricks of appropriate lengths. This design also facilitates an intuitive understanding of the acoustic principle that pitch originates from vibration frequency, which is correlated with the length of an object.

Musical fragments typically incorporate both rhythm and melody. Correspondingly, alterations in the configuration of the block track enable rolling balls to generate diverse rhythmic sequences. Continuous curves along the track cause the ball to collide rhythmically, producing drum-like beats. The curvature of the track influences the ball's speed, while its length affects the duration of the off-beat, thereby modifying the interval between notes. By constructing tracks of varying heights, children can create differences in sound intensity as the ball collides with the base, resulting in rhythmic patterns of strong and weak beats. Through adjustments to the track structure, children are able to independently construct a variety of rhythmic patterns.

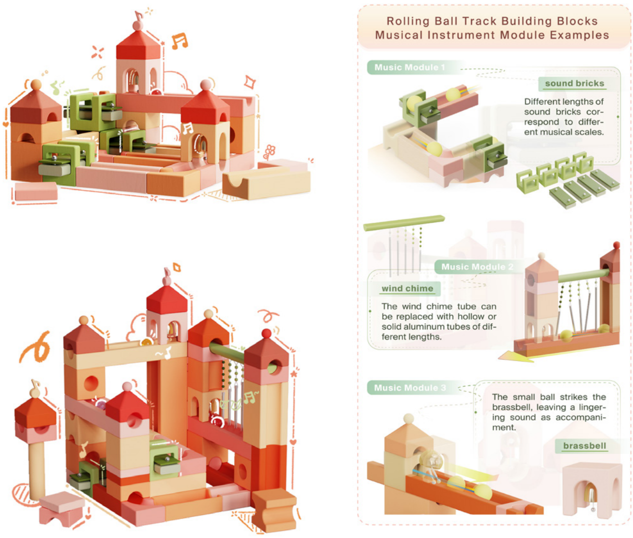

Figure 4: Music track rolling ball building blocks module.

AIoT Guided Connectivity Module

The design of the AIoT Guided Interconnection Module is informed by case studies of intelligent guidance and collaborative connectivity. Heuristic guidance preserves children's immersive play experience more effectively than directive instruction. When toys acquire social attributes, children's exploration shifts from individual to collaborative engagement, thereby exponentially increasing the learning value. In terms of physical design, the AIoT analysis and output module is primarily integrated into the accompanying light screen projector, delivering diverse interactive experiences through screen displays, projections, and voice prompts.

From a technological perspective, the building block prototypes incorporate concealed microphones that capture real-time sound sequences generated by the rolling balls. The AI system analyzes children's construction patterns and the resulting sound outputs, providing heuristic guidance through subtle light changes or brief voice prompts. For example, when a child repeatedly attempts the same track structure, the AI may intelligently suggest, "Try adding a curve block here; it might create a more soothing melody." When children place different sound bricks, it might prompt, "Would you like the collision sound to be higher in pitch? Perhaps try a shorter sound brick." This form of real-time feedback functions as an invisible playmate, preserving immersion while subtly expanding children's creative thinking. The core function of the AI is not to provide correct answers but to stimulate active exploration through heuristic questioning.

Simultaneously, the IoT module enables connectivity across multiple building block sets. Children's constructed musical tracks can trigger sound responses from peers' blocks or be uploaded to the cloud to archive musical play sequences for sharing, thereby achieving true collaborative play. In this context, sound evolves from a solitary activity into an ensemble composed with friends. Families can engage in remote collaboration via the cloud, extending music education into the social dimension. The cloud platform synchronously records children's exploration trajectories, including tonal preferences, frequently used height combinations, and creative pathways, generating a visual Music-Spatial Thinking Development Report. This approach not only renders implicit learning processes observable but also provides data-driven insights for parents and educators seeking to understand child development. AI and IoT technologies are seamlessly integrated into the building block sets, and the embedded digital intelligence ensures that children remain focused on the physical construction experience.

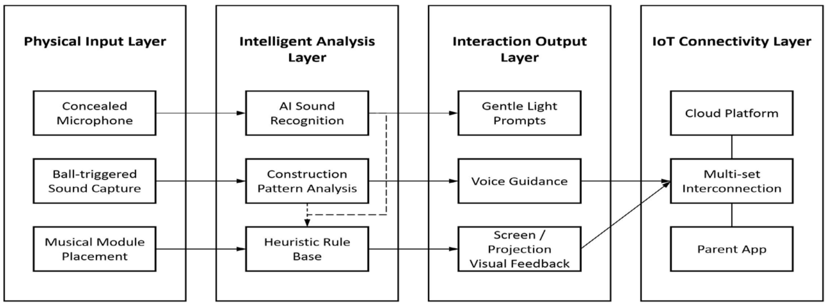

Figure 5: AIoT-guided interconnected module technology architecture.

RESEARCH LIMITATIONS

This study has several limitations. Due to practical constraints in technical implementation, the AIoT module remains at the conceptual design stage and has yet to undergo comprehensive technical validation. Although the case analysis selected three representative product categories, the limited sample size may not have encompassed all relevant design dimensions. The effectiveness of the proposed design framework therefore requires further empirical investigation to be adequately validated.

Building on these findings, future research will focus on the following directions. First, a technical prototype will be developed and subjected to small-scale user testing in order to gather authentic feedback from children. Second, the scope of case studies will be expanded to include a broader range of cutting-edge international products and research findings, thereby exploring the potential applications of AIoT technology across additional educational contexts for children.

CONCLUSION

This study demonstrates that AIoT technology not only expands the educational dimensions of traditional building block toys but also achieves an intrinsic integration between spatial construction and music generation through a physics-to-music mapping mechanism. The research contributes theoretically by introducing the first design framework for AIoT-Empowered Physical Acoustic Enlightenment Toys, thereby offering novel perspectives for the design of children's STEAM education products. Its practical value lies in providing a referenceable technical pathway for the integrated development of toy design and educational technology.

The Music Track Rolling Ball Building Blocks enable children to perceive acoustic principles through the physical process of block construction and ball collisions, while also allowing them to visualize musical structures through track design. In this way, children are able to create their own audiovisual worlds through play. It is anticipated that this design framework will inspire future research and practice, thereby contributing to the continued innovation of children's STEAM educational toys.

REFERENCES

Belpaeme, T., Kennedy, J., Ramachandran, A., Scassellati, B., & Tanaka, F. (2018). Social robots for education: A review. Science Robotics, *3*(21), eaat5954. https://doi.org/10.1126/scirobotics.aat5954

Brosterman, N. (1997). Inventing kindergarten. Harry N. Abrams.

Frischen, U., Schwarzer, G., & Degé, F. (2022). Music lessons enhance executive functions in 6- to 7-year-old children. Frontiers in Psychology, *12*, 764890. https://doi.org/10.3389/fpsyg.2021.764890

Habibi, A., Damasio, A., Ilari, B., & Sachs, M. (2022). Music training and child development: A review of recent findings from a longitudinal study. Annals of the New York Academy of Sciences, *1518*(1), 32–45. https://doi.org/10.1111/nyas.14806

Lauer, J. E., Esposito, A. G., & Bauer, P. J. (2023). Block play and spatial skills in early childhood: A longitudinal investigation. Child Development, *94*(3), 712–728. https://doi.org/10.1111/cdev.13887

McDougal, E., Silverstein, P., Treleaven, O., Jerrom, L., Gilligan-Lee, K., Gilmore, C., & Farran, E. K. (2024). Assessing the impact of LEGO® construction training on spatial and mathematical skills. Developmental Science, *27*(2), e13432. https://doi.org/10.1111/desc.13432

Miendlarzewska, E. A., & Trost, W. J. (2014). How musical training affects cognitive development: Rhythm, reward and other modulating variables. Frontiers in Neuroscience, *7*, 279. https://doi.org/10.3389/fnins.2013.00279

Pesek, M., Hirci, N., Žnideršič, K., & Marolt, M. (2024). Enhancing music rhythmic perception and performance with a VR game. Virtual Reality, *28*(2), 118. https://doi.org/10.1007/s10055-024-01014-y

Rajkumar, A. S., Kumar, S., Priya, R., Sathish, T., & Tanaka, H. (2025). Eco-Friendly IoT-based smart toy automation for interactive learning and engagement. In 2025 6th International Conference on Electronics and Sustainable Communication Systems (ICESC) (pp. 1511–1515). IEEE. https://doi.org/10.1109/ICESC65114.2025.11212517

Rath, M. O. (2003). An expressive real-time sound model of rolling [Conference paper]. *Proceedings of the 2003 International Conference on Digital Audio Effects (DAFx-03)*, London, UK. https://api.semanticscholar.org/CorpusID:17428360

Resnick, M. (2017). Lifelong Kindergarten: Cultivating Creativity through Projects, Passion, Peers, and Play. MIT Press.

Schmitt, S. A., Purpura, D. J., Duncan, R. J., Bryant, L., & Zehner, T. M. (2025). Testing block play as an effective mechanism for promoting early math, executive function, and spatial skills in preschoolers from low-income backgrounds. Early Childhood Research Quarterly, *71*, 163–173. https://doi.org/10.1016/j.ecresq.2024.12.001

Sun, L. (2025). Children's sensory interaction music enlightenment system based on sensory interaction and game elements. Entertainment Computing, *52*, 100873. https://doi.org/10.1016/j.entcom.2024.100873

Takefuji, Y. (2025). Enhancing AI education through practical IoT applications and gesture recognition. Frontiers of Digital Education, *2*(4), 29. https://doi.org/10.1007/s44366-025-0066-7

Xiao, W., & Gonçalves, A. (2025). Intelligent toys, complex questions: A literature review of artificial intelligence in children's toys and devices. Big Data & Society, *12*(2). https://doi.org/10.1177/20539517251389860