BrickSmart: Leveraging Generative AI to Support Children's Spatial Language Learning in Family Block Play

YUJIA LIU*, Tsinghua University, Beijing, China

SIYU ZHA*, Tsinghua University, Beijing, China

YUEWEN ZHANG, Tsinghua University, Beijing, China

YANJIN WANG, University of Toronto, Toronto, Ontario, Canada

YANGMING ZHANG, Wuhan University, Wuhan, Hubei, China

QI XIN, Tsinghua University, Beijing, China

LUNYIU NIE, The University of Texas at Austin, Austin, Texas, United States

CHAO ZHANG, Cornell University, New York, United States

YINGQING XU, Tsinghua University, Beijing, China

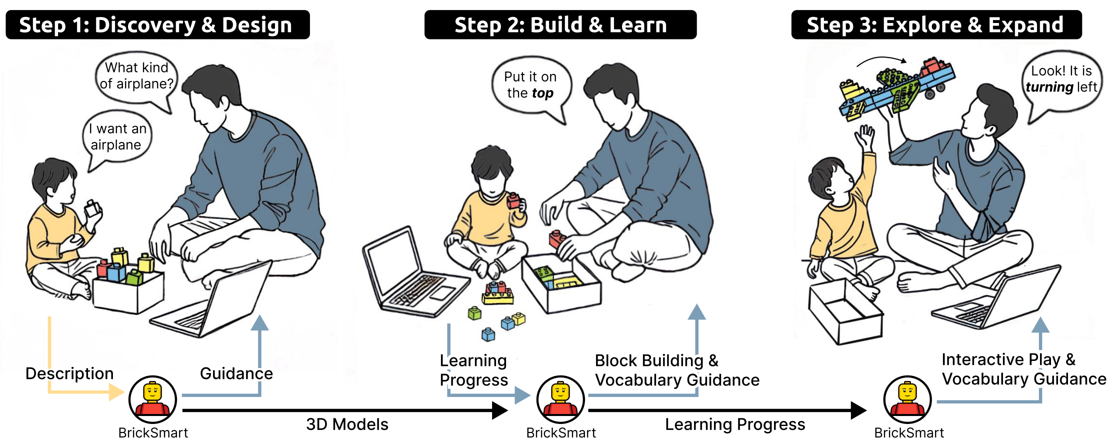

Fig. 1. BrickSmart's three-step process: Discovery & Design, Build & Learn, and Explore & Expand. Each step is designed to facilitate parent-child interactions that enhance the child's spatial language and reasoning skills through guided block design, building, and play. Key features include personalized building instructions, active learning progress tracking, and adaptive spatial vocabulary guidance. This process demonstrates a structured method for enhancing early cognitive development in children by integrating generative educational technology into traditional block play.

*Both authors contributed equally to this research.

Authors' Contact Information: Yujia Liu, Tsinghua University, Beijing, China, l-yj22@mails.tsinghua.edu.cn; Siyu Zha, Tsinghua University, Beijing, China; Yuewen Zhang, Tsinghua University, Beijing, China; Yanjin Wang, University of Toronto, Toronto, Ontario, Canada; Yangming Zhang, Wuhan University, Wuhan, Hubei, China; Qi Xin, Tsinghua University, Beijing, China; Lunyiu Nie, The University of Texas at Austin, Austin, Texas, United States; Chao Zhang, Cornell University, New York, United States; Yingqing Xu, Tsinghua University, Beijing, China.

Block-building activities are crucial for developing children's spatial reasoning and mathematical skills, yet parents often lack the expertise to guide these activities effectively. BrickSmart, a pioneering system, addresses this gap by providing spatial language guidance through a structured three-step process: Discovery & Design, Build & Learn, and Explore & Expand. This system uniquely supports parents in 1) generating personalized block-building instructions, 2) guiding parents to teach spatial language during building and interactive play, and 3) tracking children's learning progress, altogether enhancing children's engagement and cognitive development. In a comparative study involving 12 parent-child pairs for both experimental and control groups, BrickSmart demonstrated improvements in supportiveness, efficiency, and innovation, with a significant increase in children's use of spatial vocabularies during block play, thereby offering an effective framework for fostering spatial language skills in children.

CCS Concepts: • **Human-centered computing → Empirical studies in HCI**.

Additional Key Words and Phrases: Parent-child, large language model, block play

ACM Reference Format:
Yujia Liu, Siyu Zha, Yuewen Zhang, Yanjin Wang, Yangming Zhang, Qi Xin, Lunyiu Nie, Chao Zhang, and Yingqing Xu. 2025. BrickSmart: Leveraging Generative AI to Support Children's Spatial Language Learning in Family Block Play. In *Proceedings of The ACM Conference on Human Factors in Computing Systems (CHI'25).* ACM, New York, NY, USA, 26 pages. https://doi.org/XXXXXXX.XXXXXXX

1 Introduction

Spatial language, integral to daily life, describes the characteristics and relational dynamics of objects within space, such as "big/small" and "up/down" [7, 39]. The development of spatial language during childhood is pivotal as it lays the foundation for advanced spatial cognition [39], logical reasoning [15], and mathematical ability [17]. Among educational tools, block play, a prevalent family activity, is particularly noted for naturally fostering spatial language and enhancing spatial skills [15].

Vygotskian theory posits that children learn spatial language through block play more effectively with instructional guidance from an experienced partner; without it, play tends to focus purely on entertainment, foregoing educational benefits [33]. Research confirms that guided play elicits significantly more spatial language use compared to unstructured or assembly play [15]. In family settings, parents often serve as facilitators during guided block play, ideally leveraging their familiarity with the child's cognitive level and interests [10, 15, 19]. Effective facilitation requires parents to understand spatial language, recognize teachable moments, and provide structured guidance. However, many parents, particularly those from lower socioeconomic backgrounds, lack the necessary skills and knowledge to effectively foster their children's spatial language development through block play.

Human-Computer Interaction (HCI) researchers have explored numerous approaches to support parents in guiding their children's play for educational outcomes like story comprehension [44, 45], language acquisition [4], computational thinking [12], and science learning [41]. Notably, Xu et al. [40] developed a voice agent to enhance the communication between children and parents during video game-based learning of mathematical language. However, parents often find it challenging to guide block play-based spatial language learning, which demands specific skill development. To address these challenges, our study employs Generative AI (GenAI) to provide adaptive, real-time guidance, thereby enhancing parent-child interactions and promoting spatial language development during guided play.

Manuscript submitted to ACM

Recent advances in GenAI have facilitated the creation of adaptive, context-aware learning experiences that dynamically cater to a child's specific needs [13, 26, 32]. GenAI's capability to deliver personalized prompts, generate real-time feedback, and provide detailed step-by-step instructions makes it highly suitable for guided block play, where children learn through interactive, hands-on activities. Therefore, this study aims to explore the potential of GenAI in enhancing parent-guided block play, aiming to support the development of children's spatial language skills.

To address the challenges of guiding children in block play-based spatial language learning, we developed BrickSmart, a generative AI-based system designed to assist parents. BrickSmart provides four core functionalities: Systematize Scaffolding, Personalized Building Instruction Generation, Learning Progress Tracking, and Guide Suggestion Generation. These features work together to create an engaging, adaptive, and interactive learning environment that enhances both educational outcomes and parent-child interactions during block play.

To evaluate the effectiveness of BrickSmart, we conducted a comparative experiment involving a total of 24 parent-child pairs. The experimental group engaged in guided block play sessions using the BrickSmart system, while the control group did not. Results indicated significant enhancements in spatial language comprehension and usage among children who used BrickSmart, compared to their counterparts in the control group. Moreover, parents in the experimental group reported increased confidence and capability in guiding their children, attributing this to the system's adaptive guidance and real-time feedback. These findings highlight the transformative potential of generative AI in enhancing guided play as an educational experience.

In summary, the contributions of this study are as follows:

- Development of BrickSmart, aa system powered by Generative AI that improves parent-child interactions and supports spatial language learning through guided block play;
- Conducting a comparative and exploratory user study to assess the effectiveness of BrickSmart in enhancing children's spatial language skills and overall engagement;
- Provision of design insights for developing AI systems aimed at assisting parents in educational settings, offering guidelines for the future integration of AI in learning environments.

2 Related work

2.1 Spatial Language Development through Guided Block Play

Spatial language is essential for describing spatial relationships and characteristics of objects within a specific space, using terms like "below" and "behind" to intuitively segment and navigate environments [15, 39]. Early mastery of spatial language is crucial for children as it lays the foundation for advanced spatial cognition, logical reasoning, and mathematical abilities [15, 17, 39]. For young learners, understanding spatial concepts through language is challenging but fundamental for cognitive development [15, 16]. Research categorizes spatial language into eight types: dimensions, shapes, positions, orientations, quantity, deictics, attributes, and patterns [7], which provide a structured framework for developing spatial language skills.

Block play is an effective and enjoyable method for fostering spatial language development in children. It provides a hands-on environment that encourages the natural use of spatial vocabulary and cognitive skills [8, 21, 37]. Studies have shown that both children and parents use more complex spatial language during guided block play compared to free play, leading to better learning outcomes [15, 30]. In particular, guided play—an approach positioned between free play and direct instruction—has proven effective in promoting spatial language skills. It involves a facilitator, often a parent, setting clear learning objectives while allowing the child to explore actively [22, 38]. By using targeted questions, guiding statements, and heuristic prompts, facilitators can expand children's spatial vocabulary and help them build more complex structures [10, 19].

The role of adults in guided play extends beyond providing instructions; it includes offering personalized support based on the child' s existing knowledge and interests. Parents, being most familiar with their children's cognitive levels, can deliver more tailored guidance, respond to immediate needs, and provide timely feedback, creating a supportive and motivating learning environment [35, 43]. Research shows that increased use of spatial vocabulary by adults leads to a corresponding increase in spatial language use by children, highlighting the effectiveness of guided play in developing these skills [29, 33].

In summary, guided block play, supported by effective adult facilitation, provides a balanced approach to enhancing children's spatial language and cognitive development. However, many parents may lack the expertise to provide optimal guidance, presenting challenges in effectively scaffolding learning and maintaining meaningful interactions that promote spatial language skills.

2.2 Generative AI for Children's Learning

Generative AI is transforming the landscape of children's education by providing personalized, adaptive, and interactive learning experiences. Unlike traditional educational methods, AI-powered tools can dynamically generate content tailored to each child' s learning pace, needs, and interests, which is particularly effective for developing both language and cognitive skills [20, 23]. This ability to adapt in real time is crucial in creating meaningful and engaging learning environments.

AI-driven educational tools like StoryBuddy and Open Sesame leverage generative AI to create interactive narratives and exercises that adjust based on children's responses and progress [27, 45]. These tools help children expand their vocabulary and language comprehension by immersing them in scenarios that require active participation, expression, and contextual understanding. Such adaptability keeps children engaged, promotes deeper learning, and supports their cognitive development by continuously challenging them at the right level.

Moreover, the application of generative AI in educational settings aligns well with guided play, a method positioned between free play and direct instruction that has proven effective in promoting learning outcomes [22, 38]. In guided play, children benefit from having a facilitator, often a parent or educator, who provides strategic prompts and feedback to guide their learning while allowing them to explore independently [15, 30]. Generative AI enhances this process by offering adaptive guidance that evolves in response to the child's actions, effectively complementing the role of the facilitator.

Research shows that AI tools can empower parents and educators by providing real-time, adaptive suggestions that enhance their ability to support children's learning even without specialized expertise [44]. By generating context-aware prompts and feedback, generative AI helps maintain a balance between child autonomy and necessary support, which is key to effective guided play. Therefore, our study explores the potential of generative AI to enhance children's guided play by developing an AI-based system that provides adaptive, context-sensitive guidance, empowering parents and educators to support children's spatial language learning better.

2.3 AI Tools for Enhancing Parent-Led Guided Play

Traditional block-building activities often require substantial guidance from parents, who may lack the expertise to provide effective and scientifically grounded support [11]. As a result, parents often struggle to offer the kind of scaffolded learning that optimizes educational outcomes during these activities. Recent research in the HCI community has explored technological solutions, particularly AI, to provide guided play that enhances parent-child interactions [3, 34]. For example, Xu et al. developed a voice-guided game that helps children aged 4 to 7 learn mathematics by building a "math kingdom." In this game, an AI agent not only encourages children to answer questions but also provides timely feedback, enhancing their understanding of mathematical concepts. The study highlights that AI-driven parent-child games can significantly increase children's interest in learning and improve their mastery of mathematical language, which is crucial for early education [40].

Advances in natural language comprehension and question generation (e.g., [13, 26, 32]) have made it feasible to generate automated guidance that supports children's diverse learning needs while also assisting parents in their facilitative role. This enables interactive question-answering between children and AI systems that can act as collaborative partners to parents. For instance, Zhang et al. introduced StoryBuddy, an interactive AI tool designed to facilitate educational goals by engaging both AI and parents, thereby addressing the challenge of maintaining strong parent-child relationships in AI-assisted learning environments [45]. Another study developed a social robot companion that guides and motivates children during storybook reading, enhancing their exploratory learning while enabling parents to play an active role in the process [44]. Similarly, StoryCoder uses storytelling as a creative activity, allowing children to modify stories in computational thinking games, promoting creative engagement that can be enriched by parental involvement [12]. Conversational agents have also been employed to support children's literacy development [42], bilingual language acquisition [4], and science learning [41], all of which can be enhanced when parents are included in the AI-driven educational process.

AI tools offer personalized learning experiences by customizing tutorials to each child's unique needs, which can help parents provide more tailored and effective guidance. For example, Open Sesame utilizes a Target Vocabulary Extractor to identify children's vocabulary and then generates storybooks to facilitate targeted vocabulary learning through intervention methods [27]. Such AI systems can help parents provide structured support that aligns with their child's developmental needs. However, despite these advancements, few studies focus specifically on enhancing parental guidance in children's spatial language development. Therefore, this study aims to use a GAI-based agent to assist parents in guiding children during block play, promoting the development of children's spatial language skills. By leveraging AI, the study seeks to enhance the effectiveness of parental guidance in spatial language learning, empowering parents to be more confident and capable facilitators in their child's educational journey.

3 Design Goals

Based on insights from prior research on spatial language acquisition in early childhood education (refer to section 2) and reflections from our iterative design process, we have identified four key design goals (DGs) essential for designing a parent-guided, child-centered system to support spatial language learning during block-building activities.

DG1: Systematize Spatial Language Teaching Through the "What, When, How" Framework. To provide comprehensive and structured spatial language instruction, BrickSmart employs the "What, When, How" framework:

What: Identifies specific dimensions of spatial language essential for cognitive development, such as spatial relations, shapes, and orientations, which are foundational for logical thinking and mathematics [7]. **When:** Determines the optimal moments to introduce specific spatial terms during play, aligning them with key stages of the block-building activity (e.g., preparation, building, exploration) to enhance contextual learning [10]. **How:** Guides parents in using effective instructional strategies such as scaffolding, interactive questioning, and modeling language to help children understand and use spatial terms [24, 36].

DG2: Enhance Engagement through Personalized Learning Experiences. Research shows personalized learning experiences can significantly improve children's engagement and learning outcomes, particularly in early education [1, 2]. BrickSmart provides a personalized starting point where children can choose the content and projects they are interested in, allowing the system to generate corresponding block-building tasks or visual representations of the projects. This approach enables children to focus on goals that interest them, stimulating their curiosity and desire to explore. Simultaneously, the system dynamically adjusts the complexity and type of spatial language based on the child ' s developmental stage, current language proficiency, and learning progress, ensuring that the content remains challenging but not overwhelming. This combination of personalization and engagement helps children better grasp spatial concepts in a positive learning environment [2, 18].

DG3: Adaptively Track Learning Progress and Provide Feedback. Monitoring progress and providing adaptive feedback are critical to maintaining effective learning. Research indicates that continuous feedback helps reinforce learning and supports long-term retention of spatial vocabulary [6, 28, 40]. BrickSmart integrates a real-time progress-tracking system to evaluate the child ' s understanding and use of spatial language. This adaptive feature helps parents assess their child ' s progress and dynamically adjusts the learning content to ensure it aligns with the child ' s evolving needs.

DG4: Encourage Active Parent Involvement. Parental involvement is a cornerstone of successful early childhood education, especially in language acquisition [35, 43]. BrickSmart is designed to support parents by providing real-time prompts, examples, and suggestions on how to engage effectively with their child during play. This approach empowers parents to take an active role in their child's learning journey, creating a collaborative environment that fosters growth and mutual engagement.

By following these design goals, BrickSmart provides a structured yet flexible approach to enhancing spatial language development through guided play. The system ensures that both children and parents are fully supported, promoting a meaningful and impactful learning experience.

	Spatial Language Vocabularies
1. Spatial Dimensions	Big/Little, Long/Short, High/Low, Wide/Narrow, Thick/Thin, Skinny/Fat, Deep/Shallow, Full/Empty, Length, Height, Width, Depth, Volume, Capacity, Area.
2. Shapes	Circle, Square, Rectangle, Triangle, Oval, Semicircle, Polygon, Cube, Sphere, Cylinder, Cone.
3. Locations and Directions	At, To/From, On/Off, Ahead/Behind, Left/Right, In/Out, Between/Aside, Opposite, Position, Direction, Distance, Path.
4. Orientations and Transformations	Forward/Backward, Turn left/right, Upward/Downward, Flip, Rotate, Slide, Clockwise/-Counterclockwise.
5. Continuous Amount	Whole/Part, All/Half/Fraction, A lot/A little, More/Less, Same/Equal.
6. Deictics	Here, There, Where, This, That, Which.
7. Spatial Features and Properties	Line, Curve, Edge, Flat, Bent/d, Point, Acute angle, Obtuse angle, Right angle, Vertical, Parallel.
8. Pattern	Increase/Decrease, Before/After, Next/Last, First/Last, Order, Repeat, Pattern.

Table 1. Eight dimensions of spatial language and the corresponding vocabularies [7] used in the BrickSmart system.

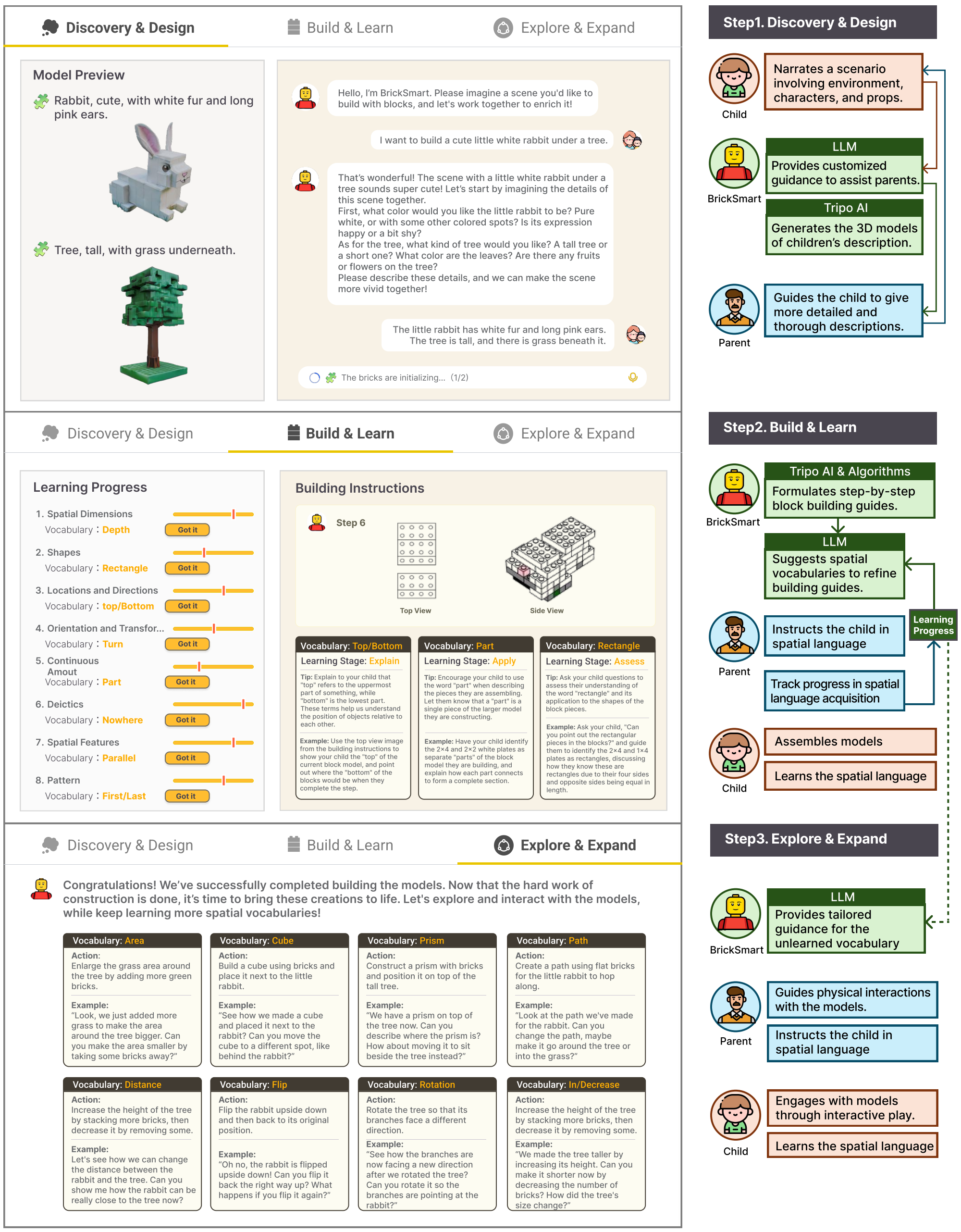

Fig. 2. Workflow of BrickSmart. **Step 1. Discover & Design:** Children describe their desired scene using voice input, and BrickSmart assists parents to help refine these ideas. The model preview appears on the left. **Step 2. Build & Learn:** Together, parents and children construct the model following the instructions of BrickSmart. Parents are advised to incorporate spatial vocabulary during the building process and track the children's learning progress. **Step 3. Explore & Expand:** The assembled models are used for interactive play, where parents introduce more spatial vocabularies to the children as guided by BrickSmart.

4 System design

To achieve our design goals, we developed BrickSmart, a GAI-based system designed to help parents guide their children in learning spatial language through block play. BrickSmart operates through three steps——Discovery & Design, Build & Learn, and Explore & Expand (as shown in fig. 2). Through these three steps, children engage deeply by building blocks they are interested in, learning spatial language systematically, and parents can track children's learning progress. BrickSmart features four core functionalities to support the workflow: Personalized Building Instruction Generation, Systematize Spatial Language Teaching, Learning Progress Tracking, Learning Progress Tracking, and Guide Suggestion Generation, each of these functionalities will be described in detail below.

4.1 [DG1] Systematize Scaffolding Spatial Language Learning.

In order to scaffold children's spatial language, we design systematised steps to guide child-parent guided block play. It is divided into three steps:Discovery & Design, Build & Learn, and Explore & Expand.

Discovery & Design: In this initial step, parents guide their children to describe the scenes they want to build, including elements like environments, characters, and props. BrickSmart takes these voice inputs from the child and generates guiding suggestions and a model preview. This process encourages children to actively participate in planning and design while providing the parent with AI-driven support to facilitate the discussion.

Build & Learn: In this step, parents and children collaboratively build the chosen scene based on the building instructions provided by BrickSmart. While constructing the model, parents use the system's guiding prompts to teach children spatial language vocabulary and concepts, with real-time updates on learning progress. BrickSmart first generates detailed building instructions and then customizes vocabulary and guidance strategies for each step based on the content and the child's current learning progress.

Explore & Expand: Once the model is built, parents and children interact with it to learn additional spatial language terms in dynamic contexts. During this stage, BrickSmart generates all remaining guiding strategies based on the completed model elements and the child's progress from the previous step. This encourages further exploration and language use, reinforcing the newly learned concepts.

Through these three steps, children engage deeply by building something they are interested in, leading to higher engagement. They learn spatial language systematically, and parents can track learning progress and adapt to different learning paces and levels of understanding. The interactive process also serves as a bridge for parent-child bonding through shared building and learning experiences.

4.2 [DG2]Personalized Building Instruction Generation

At a high-level, BrickSmart utilizes a sequence of AI-driven steps to transform initial inputs into detailed, user-friendly building guides. The process starts with the Tripo AI 3D generative diffusion model, which converts prompts into precise 3D object files in `.obj` format, ensuring accurate structure rendering. These models are then voxelized, converting the geometric data into a discrete three-dimensional matrix of 1x1x1 blocks.

An optimization algorithm [25] then refines this block matrix to improve the construction sequence. The final output is a series of layered, step-by-step building instructions, complete with top and side views for clarity. This approach highlights the effective use of AI in streamlining complex construction processes.

4.2.1 3D Model Voxelization. The task of constructing a voxelized structure using blocks can be effectively modeled as a binary programming problem. The set $\mathcal{V}$ represents the global voxel points to be covered, and set $\mathcal{B}$ includes

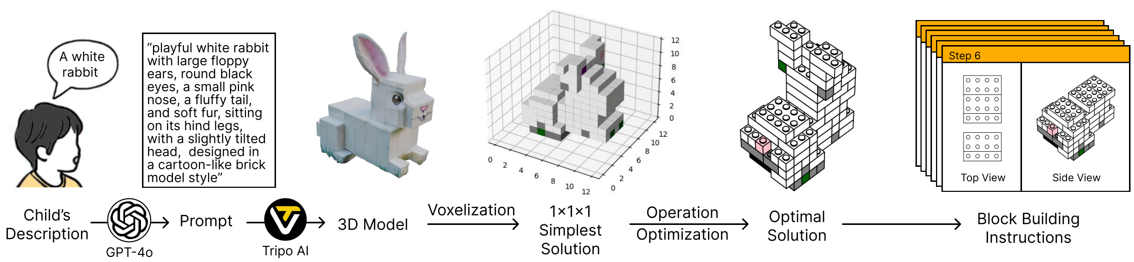

Fig. 3. Overview of the personalized building instruction generation pipeline. An LLM first revises the child's description as the prompt for 3D model generation, followed by voxelization, optimization, and the final generation of detailed block-building instructions.

all feasible brick placements. The matrix $A_{B,V}$ defines the coverage relationship between $\mathcal{B}$ and $\mathcal{V}$, with the model described by the equations:

$$\min\ z = \sum_{b\in\mathcal{B}} x_b, \qquad s.t. \sum_{b\in\mathcal{B}} a_{bv}x_b = 1, \quad \forall v \in \mathcal{V}, \qquad x_b \in 0, 1, \quad \forall b \in \mathcal{B} \tag{1}$$

The objective function (1) minimizes the number of bricks used, while constraint ensures every voxel v is covered precisely once. However, the binary programming model is limited in practical applications due to its computational intensity, lack of real-time flexibility, and potential discrepancies between optimal coverage and structural stability, as shown in Figure 4 (a), where the different structural robustness is due to staggered gaps.

4.2.2 Brick Placement Optimization. To address computational challenges and enhance interactivity, we implemented a three-stage heuristic algorithm based on the 'matheuristic' approach [25]. The first stage involves decomposing the 3D object into $1 \times X$ strips processed layer by layer, as shown in Figure 4 (b). This reduction from three to two dimensions significantly speeds up computations, facilitating real-time instruction generation.

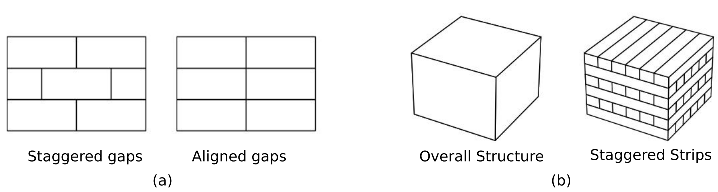

Fig. 4. (a) Comparative illustration of brick placement stability: staggered versus aligned gaps. (b) Decomposition of the object into orthogonal staggered strips, layer by layer.

The second stage uses a 1D-heuristic algorithm to segment co-planar strips into the minimum number of bricks, incorporating gap information G to optimize placement:

$$F_{cost}(b, L) = M(l_b, l_0) + N(L, l_0) + D(b, G) + e \tag{2}$$

Here, b represents the brick selected in the current step, with l_b and r_b denoting the brick's length and starting coordinate, respectively. The length l_0 corresponds to the longest brick available, and L represents the remaining length of the target strip. The multiplicity function:

Algorithm 1 Brick Placement Optimization

Input: A matrix of overall voxelized structure *voxel_matrix* and a list of brick size *brick_list*
Output: A list of bricks to make up the structure
strip_list, gap_list ← SEGMENT(*voxel_matrix*)
build_list ← NULL
for *strip* in *strip_list* **do**
  L ← length of *strip*
  for *brick* in *brick_list* **do**
    cost ← F_{cost}(*brick*, *L*)
    find the lowest *cost* and corresponding *brick*
    update *build_list*
  end for
    update *gap_list*
end for
build_list ← MERGE(*build_list*)
return *build_list*

$$M(l_b, l_0) = \frac{l_0}{l_b} \tag{3}$$

aims to minimize the frequent use of smaller bricks. The remainder counting function:

$$N(L, l_0) = \begin{cases} \left\lceil \frac{L-\rho}{l_0} \right\rceil + LpInt(\rho), & L \geq \rho \\ LpInt(L), & L < \rho \end{cases} \tag{4}$$

where ρ determines when to engage integer linear programming for the remaining part. The optimization is described as follows, similar to the binary programming model:

$$min\ N = \sum_{b^* \in \mathcal{B}^*} x_{b^*}, \qquad s.t. \sum_{b^* \in \mathcal{B}^*} l_{b^*} x_{b^*} = L, \qquad x_{b^*} \in \mathbb{Z}^+, \quad \forall b^* \in \mathcal{B}^* \tag{5}$$

The border-gap evaluation function is defined as:

$$D(b, G) = \gamma_1 \exp(-\gamma_2 d_{b,G}) \tag{6}$$

where $d_{b,G}$ is the closest vertical distance between the borders of brick b and the recorded gaps G, and γ_1, γ_2 are exponential coefficients. The term e represents a random perturbation, introducing variability into the algorithm to avoid local minima and enhance solution diversity. This holistic approach addresses both efficiency in brick usage and aesthetic considerations by considering gap placements within the overall structure.

In the final stage, adjacent bricks are merged to form larger units, reducing the number of components and simplifying the building instructions. For example, two adjacent 1 × 4 bricks are combined into a single 2 × 4 brick. The segmented and merged results are then compiled into a comprehensive set of building instructions, as illustrated in Figure 2.

Algorithm 1 provides a detailed description of the entire three-stage procedure. This ensures that the building process is both efficient and user-friendly, addressing the challenges of computational complexity while adhering to practical construction techniques.

4.3 [DG3] Learning Progress Tracking

We developed the Learning Progress Tracking feature to cater to each child's unique learning pace and provide parents with a comprehensive understanding of their child's progress. This functionality tracks the child's mastery of spatial language across eight key dimensions: Spatial Dimensions, Shapes, Locations and Directions, Orientations and Transformations, Continuous Amount, Deictics, Spatial Features and Properties, and Pattern (as detailed in Table 1).

For each dimension, the system displays a visual progress tracker for specific vocabulary terms, such as "Big/Little," "Circle," "Left/Right," "Rotate," and "Increase/Decrease." Parents can easily see which terms have been mastered and which require further practice. This level of detail allows parents to focus on areas that need reinforcement, ensuring a balanced and comprehensive development of their child ' s spatial language skills.

Moreover, by providing real-time updates on progress, the system helps parents make informed decisions about adjusting the learning plan and tailoring future activities. This dynamic, data-driven approach enables personalized learning paths that adapt to different learning speeds and comprehension levels. It empowers parents to actively engage in their child's educational journey and optimize the learning experience.

4.4 [DG4]Structured Parental Guidance and Suggestion Generate

During and after the block-building process, BrickSmart utilizes GPT-4 to generate structured guidance and suggestions that help parents facilitate their child's learning experience. These AI-generated prompts are designed to provide clear, context-sensitive instructions that support both spatial language development and the building activity itself. The prompts guide parents in introducing relevant vocabulary, asking questions to encourage spatial reasoning, and offering constructive feedback to maintain engagement and learning momentum. An example of these prompts is provided in Table 6. This approach ensures that parents are equipped with the right tools to create a rich, interactive, and educational block play experience.

5 User Study

We conducted a comparative study to evaluate the effectiveness and usability of BrickSmart in supporting spatial language development during block play with children aged 6 to 8 years. This study aimed to understand how BrickSmart's guided play approach influences children's spatial vocabulary acquisition, engagement, and overall learning experience. We hypothesize that the personalized spatial language guidance provided by BrickSmart will be perceived as a valuable tool by both children and parents, enhancing children's spatial language skills during block-building activities.

5.1 Participants

We conducted a user study to evaluate the effectiveness of the BrickSmart system in enhancing spatial language development among children aged 6 to 8 years. A total of 24 parent-child pairs participated in the study, recruited through local community centers and online parent groups. The participants were divided into two groups: an experimental group (12 pairs) using the BrickSmart system and a control group (12 pairs) using traditional block-building methods without system guidance.

To ensure a balanced comparison, children in both groups were matched based on age, gender, and initial spatial language abilities, which were assessed through a pre-screening test. All parents provided informed consent for participation, and ethical approval for the study was obtained from the university's Institutional Review Board (IRB). The

Steps	Goal	System Prompts
Step 1	Generate descriptions	Your task is to break down the story and scenes described by the child into several describable 3D objects, and output them sequentially into a structured string list object_list. For example, if the child says "A monkey with big eyes is climbing a tree," the output string list should be: object_list= Monkey, big eyes, action is climbing a tree, Tree. Please only break down and output based on the child's description without adding extra information or your own ideas. If the description lacks sufficient detail, output an empty string list object_list. You already have an initial list: {object_list}. Based on new descriptions or existing content, you need to refine and complete the entries in the list. Here is the conversation history so far: {chat_history}.
	Generate prompts for 3D models	You are an expert at crafting text prompts for generating 3D models, specializing in transforming children's imaginative words and narrated diaries into delightful, cartoon-style 3D models. Your prompts should focus on describing a single object rather than a scene, ensuring that the description is suitable for conversion into LEGO models. When rewriting user input, consider the following: The generated 3D model should avoid unnecessary details, should have a suitable center of gravity, and should clearly and independently represent a single entity. The appearance can be slightly enhanced to appeal to the aesthetic preferences of 6-8-year-old children. Ensure the model is suitable for construction using only basic LEGO bricks, emphasizing square and rectangular forms. Your output should be a detailed sentence or a list of descriptive words separated by commas, written in English.
Step 2	Understanding tutorial images	Based on the provided LEGO assembly tutorial image, thoroughly analyze and describe the current building task, outputting a string parameter instruction: (1) Describe the overall structure and design features of the LEGO model shown in the image. (2) Identify and describe the types of LEGO pieces and their colors appearing in the image. (3) Outline the assembly steps shown in the image, including any special assembly techniques or details that require special attention. (4) Use professional LEGO terminology to enhance the accuracy and professionalism of the description. (5) Ensure the text description accurately reflects the image content, with clear, professional, and detailed language to facilitate understanding of the assembly process.
	Vocabulary selection	You are a spatial language teacher, responsible for selecting spatial language tasks based on the task and student's proficiency. Use the following information: LEGO assembly tutorial description: {instruction}. Eight spatial language dimensions: ... User's current spatial language proficiency: {understand_level}, representing progress in each dimension (as a percentage). Select 3 suitable spatial language categories, focusing on those relevant to the LEGO task and where proficiency is lower. Output a list spatial_list with 3 elements, each from 0 to 7, representing category indices.
	Guidance generation	You are a family guide helping parents improve their child's spatial language. Based on the current LEGO tutorial, generate real-time prompts for parents to teach these words: Word: word_1, Stage: {stage_1}; Word: {word_2}, Stage: {stage_2}; Word: {word_3}, Stage: {stage_3}. The current tutorial is: {instruction}, showing a top view (current step) and a whole view (completed so far). Use the building task to teach the words, matching each to its learning stage. Example format: Word: Circle, Stage: Noun Explanation. Prompt: Explain that a circle has no corners, and all points on the edge are equally distant from the center. Example: Point out circular bricks or designs to help understanding.
Step 3	Guidance generation	You are an assistant helping parents improve their child's spatial language skills. They've built LEGO models: {objects}. Your task is to guide parents to move these models and have the child describe the actions to enhance spatial understanding. Output format: Vocabulary: The word to learn. Movement Example: Specific movement of the object. Parent Prompt: Example guidance for parents. Example: Vocabulary: Left/Right. Movement Example: Move the figure forward, then turn left. Parent Prompt: "Look, the figure turns left. Can you make it turn right?". For the num_words keywords: {keywords}, provide suggestions.

Table 2. The detailed prompts BrickSmart employs across three steps. Each step is tailored to enhance children's spatial reasoning and language skills through structured interactions and tasks.

study also ensured the safety and comfort of all child participants by having a researcher present during all sessions to monitor their well-being and provide support as needed.

Fig. 5. Children's Block Designs. A collection of diverse and creative block constructions by children using the BrickSmart system.

5.2 Study Procedure and Protocol

The study employed a between-subjects experimental design, where each parent-child pair participated in one of two conditions: the experimental condition using the BrickSmart system or the control condition using traditional block-building activities. The study procedure was structured into three phases: pre-test (10 minutes), intervention (30 minutes), and post-test (20 minutes), lasting approximately one hour for each pair. The following steps outline the procedure for both groups:

5.2.1 Experimental Group: **Pre-Test (10 minutes):** Children completed a spatial language assessment (shown in Appendix **??**) describing illustrated scenes featuring various spatial relationships and objects. This served as a baseline for evaluating changes in spatial language proficiency.

Intervention (30 minutes):

- **Step 1:** Children were prompted to articulate scenes, characters, and props they wanted to build using blocks. This initial step encouraged creativity and set the stage for the building activity.
- **Step 2:** Based on the child's input, BrickSmart generated custom building models and provided step-by-step tutorials tailored to the child's ideas. Parents were guided by BrickSmart to use spatial language prompts during the building process, helping the child understand spatial terms (e.g., "above," "next to," "between").
- **Step 3:** After construction, children engaged in interactive play with their models, with BrickSmart offering additional spatial language prompts during movement and play. This helped reinforce vocabulary through context-driven interactions.

User	Group	Gender	Parent	Age	Prior Frequency of Brick-playing	Prior Habits of Brick-playing
P1	Experimental	Boy	Father	8	Often	With Parent
P2	Control	Boy	Mother	6	Occasionally	Alone
P3	Experimental	Boy	Mother	6	Often	Alone
P4	Control	Girl	Mother	7	Often	Alone
P5	Experimental	Boy	Mother	6	Occasionally	With Parent
P6	Control	Boy	Mother	6	Often	With Parent
P7	Control	Boy	Mother	6	Often	With Parent
P8	Experimental	Boy	Mother	6	Often	Alone
P9	Experimental	Boy	Father	6	Often	With Parent
P10	Experimental	Boy	Mother	6	Often	Alone
P11	Control	Girl	Grandmother	6	Often	With Parent
P12	Experimental	Girl	Mother	6	Occasionally	With Parent
P13	Control	Boy	Mother	6	Often	With Parent
P14	Experimental	Boy	Mother	6	Occasionally	With Parent
P15	Experimental	Girl	Mother	6	Occasionally	With Parent
P16	Control	Girl	Father	6	Occasionally	With Parent
P17	Experimental	Boy	Mother	6	Often	Alone
P18	Control	Girl	Mother	6	Occasionally	With Parent
P19	Control	Girl	Mother	7	Occasionally	With Parent
P20	Experimental	Boy	Father	7	Occasionally	With Parent
P21	Control	Boy	Mother	8	Often	Alone
P22	Experimental	Boy	Mother	7	Occasionally	Alone
P23	Control	Boy	Mother	6	Often	Alone
P24	Control	Girl	Mother	6	Often	Alone

Table 3. Participant Overview in BrickSmart Study. This table lists the demographics of children in both experimental and control groups, detailing their age, gender, parent involvement, and the frequency of their prior engagement with brick-based activities.

- **Step 4:** To conclude, children described their creations in a one-minute narrative, further reinforcing their use of newly learned spatial language terms.

Post-Test (20 minutes): The spatial language assessment was repeated using the same illustrated scenes as in the pre-test to measure improvements in spatial language development.

5.2.2 Control Group: **Pre-Test (10 minutes):** The control group followed the same pre-test procedure as the experimental group to establish a baseline for spatial language abilities.

Intervention (30 minutes):

- **Step 1:** Parents were briefed on the importance of using spatial language during play but were not provided with specific guidance or system-generated prompts.
- **Step 2:** Children selected from a set of pre-generated tutorials (e.g., building a rabbit, house, or tree) and followed the instructions with their parents' assistance. This process aimed to simulate natural play without the tailored support offered by BrickSmart.
- **Step 3:** Similar to the experimental group, children described their creations in a one-minute narrative, which aimed to encourage the use of any spatial language learned during the session.

Post-Test (20 minutes): A post-test similar to that of the experimental group was conducted to reassess the child's spatial language abilities using the same illustrated scenes.

5.3 Data Collection and Analysis

5.3.1 Pre-test and Post-test. Children's spatial language proficiency was evaluated using a pre-test and post-test designed with comparable difficulty. Each test included 24 questions: 16 fill-in-the-blank tasks and 8 comprehension tasks, with a total possible score of 48 points. These tests aimed to measure improvements in spatial language use after the intervention.

5.3.2 Parent Feedback and Usability Measures. After the study, parents provided feedback on the system by completing three questionnaires: the User Experience Questionnaire (UEQ), which measured user satisfaction across various dimensions; Additionally, parents participated in interviews to offer qualitative insights into the system's effectiveness and usability.

5.3.3 Children's Engagement and Fun. The Fun Toolkit [31] was used to measure children' s engagement and enjoyment, focusing on Endurability, Engagement, and Expectations.

5.3.4 Video Transcription and Coding. Video recordings of parent-child interactions during the study were transcribed verbatim to capture the use of spatial language. A coding scheme was developed to identify and categorize spatial language terms and two independent coders analyzed the transcripts. The frequency and density of spatial language usage were calculated to understand how effectively spatial language was integrated into the dialogues, highlighting the impact of the BrickSmart system on language development during block-building activities.

Quantitative data (test scores, UEQ, SUS) were analyzed using t-tests and ANOVA to compare between groups' spatial language improvement and system usability. Qualitative data from interviews and narratives were thematically analyzed to identify key insights on user experience and system effectiveness.

5.4 Results

5.4.1 Evaluation of System Usability. Statistical analysis using independent-sample t-test and Mann-Whitney U-test confirmed significant differences in several dimensions compared to the control group (as shown in Fig 6). Participants rated BrickSmart significantly better in terms of supportiveness ($t = 2.22, p = 0.037$), inventiveness ($t = 2.55, p = 0.018$), leading edge ($t = 2.08, p = 0.049$), and hedonic quality($t = 2.04, p = 0.049$)

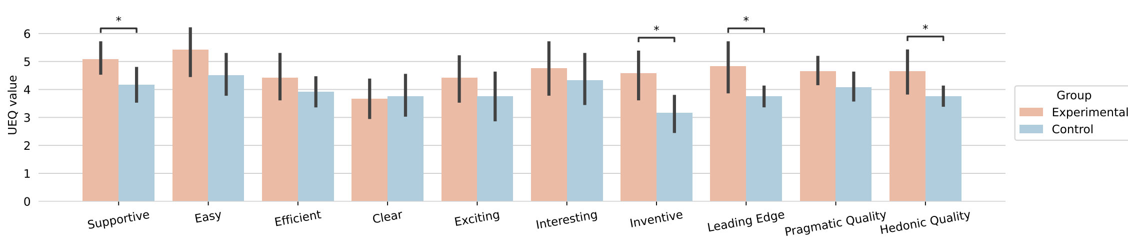

Fig. 6. Comparison of UEQ metrics between experimental and control groups. Error bars represent 95% confidence intervals (CI). * stands for $p < 0.05$ and ** stands for $p < 0.01$. The same annotation applies to the rest of the paper.

The System Usability Scale (SUS) results further indicate that BrickSmart outperforms the control condition regarding usability. The experimental group, using BrickSmart achieved a SUS score of 71.46. This suggests that users found BrickSmart to be relatively intuitive and user-friendly, reinforcing its effectiveness in enhancing spatial language development through guided block play. P8' s parent noted that the system allowed them to understand their child's cognitive progress better, while the child developed a preliminary understanding of spatial concepts through play. P3's mother highlighted that integrating vocabulary learning with LEGO building provided valuable educational insights, suggesting that more detailed guidance could further enhance the learning experience. She also mentioned that building with specific learning goals introduced new focus areas, indicating that structured play could lead to more targeted educational outcomes. Overall, these results suggest that BrickSmart not only supports children's learning of spatial language but also provides parents with a framework to better guide and understand their child's learning process, potentially leading to richer educational experiences.

5.4.2 Children's Spatial Language Skill Improvement and Performance. We evaluated the impact of the BrickSmart system on children's spatial language skills using three approaches: pre- and post-test assessments of children's knowledge, parental evaluations of their children's progress, and analysis of video transcriptions to measure the density and frequency of spatial language use in parent-child dialogues. These combined data provide a comprehensive understanding of how the system enhances spatial language development.

Student Test Results on Spatial Language Skill As shown in Fig. 7, it presents the Spatial Language Questionnaire results, which measured children ' s spatial language skills before and after the study for both groups. Statistical analysis shows that both groups significantly improved in spatial language skills from pre-test to post-test (Experimental: $t = 10.24, p < 0.001$. Control: $U = 1.0, p = 0.004$). However, the experimental group demonstrated a significantly greater improvement compared to the control group. The post-test scores of the experimental group were markedly higher($U = 127.0, p = 0.002$), indicating that the BrickSmart system was more effective in enhancing children's spatial language abilities than the traditional approach.

Parental Assessment of Children's Spatial Language Improvement Figure 7 shows the before-and-after evaluations from parents regarding their children's spatial language development. The results reveal that both groups exhibited an increase in perceived spatial language skills after the study ($all p < 0.05$), but the increase was more pronounced in the experimental group. Parents in the experimental group reported significantly higher improvements in their children's spatial language cognition compared to the control group($U = 117.5, p = 0.007$). This suggests that the BrickSmart system not only enhances children's spatial language skills but also leads to noticeable improvements as perceived by their parents.

The parents' interview feedback further supports these quantitative findings. P8's mother said, *"Spatial language is a crucial part of cognitive training for children, primarily acquired through natural learning and school materials. During this activity, I noticed my child learning about different categories of spatial language and gaining new educational insights."* Similarly, P3 ' s parent noted that the experiment helped their child recognize the importance of spatial language, but also mentioned that some new terms introduced during the study were challenging to encounter in regular block-building activities.

Analysis of Spatial Language Usage from Transcribed Dialogues Based on the transcription and coding of videos recorded during the study sessions, we analyzed the density and frequency of spatial language used in dialogues between children and parents. As shown in Figure 8, the results show a comparison of the frequency of different categories of spatial language vocabularies used in parent-child dialogues between the experimental group and the baseline group. The experimental group shows a notably higher frequency of spatial language terms across several categories. For example, in the spatial dimensions category, the experimental group used these terms over 200 times, while the baseline group used them less than 100 times. Similarly, the locations and directions category shows a significant

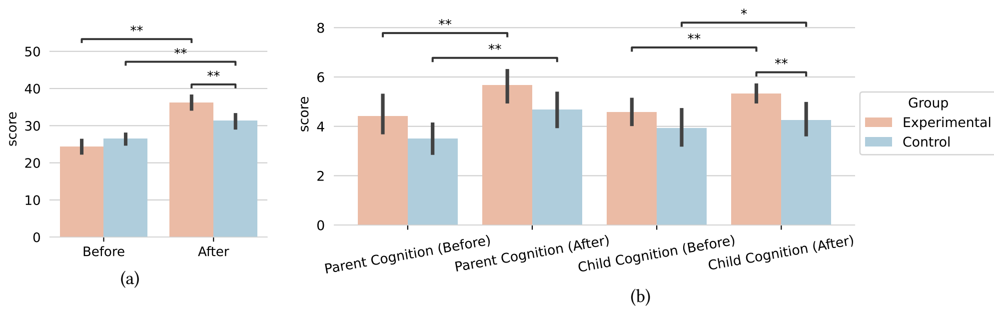

Fig. 7. Cognitive Score Comparisons: (a) Changes in overall scores before and after the intervention for experimental and control groups. (b) Detailed before and after scores for parent and child cognition in both groups, with significant differences marked.

increase in the experimental group, with terms used nearly 300 times compared to around 150 times in the baseline group.

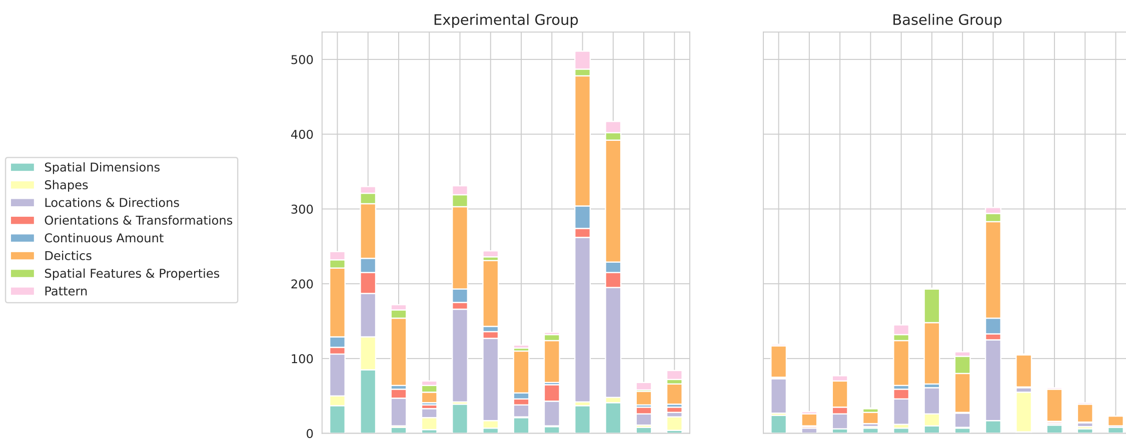

Fig. 8. Comparison of category frequencies in spatial language vocabularies' presence. The experimental group with BrickSmart demonstrates higher overall frequency and more comprehensive coverage across dimensions.

The orientations and transformations category also saw a marked increase in the experimental group, reflecting a more comprehensive use of terms describing spatial orientation changes. These were used nearly 150 times compared to less than 50 in the baseline. Additionally, terms related to deictics (e.g., "here," "there") and continuous amount (e.g., "more," "less,") were more frequently used by the experimental group, highlighting their engagement with more complex spatial concepts. Qualitative feedback from parents supports these findings. P15 noted that their child became more accurate in describing spatial orientations. P22 mentioned, *"I noticed that when positioning objects during spatial configuration, along with hands-on manipulation, my child' s spatial abilities seemed to improve."* Similarly, P14 observed an enhancement in their child's spatial language skills, stating, *"I could see a deeper understanding of spatial concepts when my child compared overall images with bird's-eye views."*

These findings suggest that the BrickSmart system promoted a richer and more varied use of spatial language, providing children with more opportunities to practice and internalize these concepts during block-building activities. This increased diversity and frequency in spatial language usage in the experimental group compared to the baseline group illustrates the system's effectiveness in enhancing spatial language development through guided interaction.

5.4.3 Children and Parent Engagement Across the System. Figure 9 illustrates the results of the engagement and fun assessments for children across different dimensions: expectation, engagement, and endurability. The results show that while both groups had similar expectations before the study, the experimental group reported significantly higher expectations after using BrickSmart ($U = 28.5, p = 0.008$). Additionally, engagement levels were notably higher in the experimental group ($t = 3.46, p = 0.002$), indicating that the interactive features and guided prompts in BrickSmart helped sustain children's interest and focus throughout the activities. In terms of endurability, or the desire to continue using the system, the experimental group also scored higher than the control group ($t = 3.55, p = 0.002$).

Parents in the control group frequently mentioned feelings of achievement (P11, P16, P18, P19, P23), while those in the experimental group emphasized different aspects. For example, P5 noted that BrickSmart aligned well with their child's interests, and P22 mentioned that their child was more engaged because the activities related to recent experiences and their interest in LEGO. P15 highlighted that the system's tutorials were varied and flexible, unlike rigid, rule-based instructions, allowing for greater creativity.

An interesting observation was that while both groups had elements of creative freedom—such as choosing different colors when specific ones were unavailable—the experimental group demonstrated a higher degree of flexibility. Children in the experimental group were more likely to alter shapes or incorporate additional LEGO pieces, suggesting that the initial guided storytelling in Step 1 might have sparked their creative instincts.

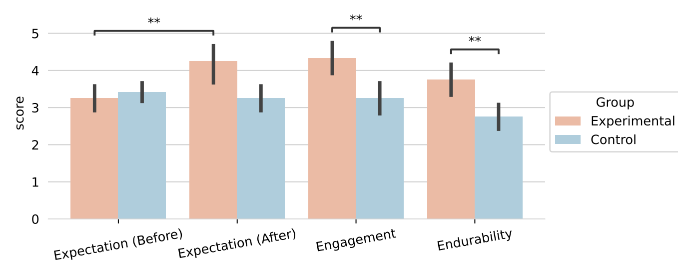

Fig. 9. The engagement and fun assessment results. Notable increases in post-use expectations and engagement were observed in the experimental group, as well as a marked increase in endurability.

Parent feedback also highlighted the system's role in enhancing parent-child interaction. While some parents in the control group, like P21, felt that good instructions reduced the need for parental involvement, others, such as P23, observed that children mostly worked independently with minimal parental intervention. In contrast, parents in the experimental group, such as P22, noted frequent interactions with their children, offering immediate encouragement when they encountered difficulties. P8 pointed out a balanced experience, noting that while the system's guidance was helpful, having specific goals also created a sense of urgency.

Most parents from both groups acknowledged that collaboration during the activities helped strengthen their bond with their children. For instance, P7 (from the experimental group) mentioned that having a clear guide and a goal that interested the child made collaboration easier and increased the child's participation. This aligns with our earlier claim that the system can serve as a ”bridge” for parent-child communication, enhancing engagement and cooperation during guided play.

Overall, these findings suggest that BrickSmart not only enhances children’s engagement and desire to learn but also fosters meaningful parent-child interactions, making it an effective tool for guided play that supports both educational and relational outcomes.

6 Discussion

This paper explores the potential of BrickSmart, an AI-driven system, in enhancing children’s spatial language development through guided block play, contributing to the growing body of research on AI-supported educational tools and human-AI collaboration in learning contexts. There are varying perspectives on the role of AI in augmenting parent-child interactions for educational purposes. Some studies suggest that AI can effectively scaffold learning by providing personalized and context-sensitive feedback, thereby enhancing the educational experience[5, 9, 14]. However, others caution against relying solely on AI systems, emphasizing the importance of human facilitation and the need for tools that support rather than replace parental involvement in children’s learning[3, 34, 45]. Our work aligns with the latter perspective by presenting BrickSmart as a system that integrates AI support while enhancing the role of parents as active facilitators during guided play. In this section, we discuss the insights gained from designing and evaluating BrickSmart, the implications for future AI-based educational tools, the limitations of our current study, and potential avenues for future research.

6.1 Personalized Building Instructions as Adaptive Scaffolding

The Personalized Building Instruction Generation in BrickSmart serves as adaptive scaffolding, enabling tailored guidance based on each child’s unique learning pace and needs. This aligns with prior research suggesting that personalized, responsive learning environments can significantly enhance engagement and learning outcomes [40, 45]. BrickSmart dynamically adjusts building instructions to the child’s preferences and current level of understanding, making the block-building process more engaging and accessible. This personalized approach ensures that children are neither overwhelmed by complexity nor disengaged due to lack of challenge. The adaptability of AI in this context demonstrates its potential to provide just-in-time scaffolding that supports children’s learning while respecting their autonomy. Future educational tools should continue to explore how adaptive AI can cater to diverse learner profiles, encouraging exploration and creativity within structured learning environments.

6.2 Guiding Suggestions for Parents to Enhance Engagement

BrickSmart also incorporates Guide Suggestion Generation, which provides real-time, context-sensitive prompts for parents to facilitate deeper engagement during play. This feature addresses a critical need in guided learning: empowering parents with the tools to effectively scaffold their child’s language development without needing extensive pedagogical expertise [3, 34]. By offering suggestions that guide parents on how to introduce spatial terms or ask thought-provoking questions, BrickSmart supports richer parent-child interactions. This aligns with findings emphasising the importance of active adult participation in enhancing children’s learning experiences [44]. However, balancing AI suggestions with parental autonomy is crucial to ensure the interactions remain natural and meaningful. Future systems should refine this balance, perhaps through customizable guidance options that allow parents to tailor the AI’s input based on their comfort and the child’s responsiveness.

6.3 Limitation and Future Work

While BrickSmart shows promise in enhancing spatial language learning, several limitations must be acknowledged. First, the study's sample size and diversity may limit the generalizability of the findings. Future studies should include a broader demographic range to assess the system's effectiveness across different contexts and cultures. Second, the study focused on short-term learning outcomes, leaving the long-term impact of using BrickSmart on children's spatial language development an open question. Future research should consider longitudinal studies to evaluate sustained learning benefits. Third, while BrickSmart provides effective AI-generated suggestions, it may not fully replicate the nuanced guidance a skilled human facilitator can offer. Enhancing AI's ability to provide more contextually aware and emotionally intelligent feedback could be a potential area for development. Lastly, integrating multimodal inputs, such as voice or gesture recognition, could enrich the interaction experience, making it more natural and engaging for parents and children.

Overall, our study highlights the potential of AI-driven systems like BrickSmart in supporting block-guided play and enhancing children's language development. Future research should continue to refine these tools, ensuring they complement human facilitation and create enriched, adaptive learning experiences for children.

7 Conclusion

This study introduced BrickSmart, a GenAI-driven system designed to support parents in enhancing children's spatial language development through guided block play. By leveraging natural language processing and interactive prompts, BrickSmart facilitates meaningful parent-child interactions, creating a more effective learning environment for spatial language acquisition. Our findings from both controlled and exploratory studies demonstrate that BrickSmart significantly improves children's spatial language skills and engagement compared to traditional unguided play. Additionally, the study highlights the importance of integrating GenAI in educational contexts to empower parents as facilitators of learning. The insights gained from the design and evaluation of BrickSmart provide valuable guidelines for developing future GenAI tools that enhance educational outcomes by supporting guided play and other interactive learning methods. Future work will explore expanding the system' s capabilities and examining its application in broader educational contexts.

References

[1] Ohoud Almousa and Sharifa Alghowinem. 2023. Conceptualization and development of an autonomous and personalized early literacy content and robot tutor behavior for preschool children. *User Modeling and User-Adapted Interaction* 33, 2 (2023), 261–291.

[2] Hee Jin Bang, Linlin Li, and Kylie Flynn. 2023. Efficacy of an adaptive game-based math learning app to support personalized learning and improve early elementary school students' learning. *Early Childhood Education Journal* 51, 4 (2023), 717–732.

[3] Ilene R Berson, Michael J Berson, Candice McKinnon, Deeksha Aradhya, May Alyaeesh, Wenwei Luo, and Ben Rydal Shapiro. 2023. An exploration of robot programming as a foundation for spatial reasoning and computational thinking in preschoolers' guided play. *Early Childhood Research Quarterly* 65 (2023), 57–67.

[4] Neelma Bhatti, Timothy L Stelter, Scott McCrickard, and Aisling Kelliher. 2021. Conversational user interfaces as assistive interlocutors for young children' s bilingual language acquisition. In *Proceedings of the 2021 ACM International Conference on Interactive Media Experiences*. 208–211.

[5] Muhammad Roil Bilad, Lalu Nurul Yaqin, and Siti Zubaidah. 2023. Recent progress in the use of artificial intelligence tools in education. *Jurnal Penelitian dan Pengkajian Ilmu Pendidikan: e-Saintika* 7, 3 (2023), 279–314.

[6] Andrew Thomas Bimba, Norisma Idris, Ahmed Al-Hunaiyyan, Rohana Binti Mahmud, and Nor Liyana Bt Mohd Shuib. 2017. Adaptive feedback in computer-based learning environments: a review. *Adaptive Behavior* 25, 5 (2017), 217–234.

[7] Joanna Cannon, Susan Levine, and Janellen Huttenlocher. 2007. A system for analyzing children and caregivers' language about space in structured and unstructured contexts. *Spatial Intelligence and Learning Center (SILC) technical report* (2007).

[8] Beth M Casey, Nicole Andrews, Holly Schindler, Joanne E Kersh, Alexandra Samper, and Juanita Copley. 2008. The development of spatial skills through interventions involving block building activities. *Cognition and Instruction* 26, 3 (2008), 269–309.

[9] Xieling Chen, Di Zou, Haoran Xie, and Gary Cheng. 2021. Twenty years of personalized language learning. *Educational Technology & Society* 24, 1 (2021), 205–222.

[10] Lynn E Cohen and Janet Emmons. 2017. Block play: spatial language with preschool and school-aged children. *Early Child Development and Care* 187, 5-6 (2017), 967–977.

[11] Lauren DeLaCruz. 2021. A Toy for Thoughtful Parents: The Rhetorical Building Blocks of LEGO's Reputation. *The Marketing of Children's Toys: Critical Perspectives on Children's Consumer Culture* (2021), 165–183.

[12] Griffin Dietz, Jimmy K Le, Nadin Tamer, Jenny Han, Hyowon Gweon, Elizabeth L Murnane, and James A Landay. 2021. Storycoder: Teaching computational thinking concepts through storytelling in a voice-guided app for children. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–15.

[13] Xinya Du, Junru Shao, and Claire Cardie. 2017. Learning to ask: Neural question generation for reading comprehension. *arXiv preprint arXiv:1705.00106* (2017).

[14] Jorge Fernández-Herrero et al. 2024. Evaluating Recent Advances in Affective Intelligent Tutoring Systems: A Scoping Review of Educational Impacts and Future Prospects. (2024).

[15] Katrina Ferrara, Kathy Hirsh-Pasek, Nora S Newcombe, Roberta Michnick Golinkoff, and Wendy Shallcross Lam. 2011. Block talk: Spatial language during block play. *Mind, Brain, and Education* 5, 3 (2011), 143–151.

[16] Dedre Gentner, Asli Özyürek, Özge Gürcanli, and Susan Goldin-Meadow. 2013. Spatial language facilitates spatial cognition: Evidence from children who lack language input. *Cognition* 127, 3 (2013), 318–330.

[17] Carrie Georges, Véronique Cornu, and Christine Schiltz. 2023. The importance of spatial language for early numerical development in preschool: Going beyond verbal number skills. *Plos one* 18, 9 (2023), e0292291.

[18] Peggy Grant and Dale Basye. 2014. *Personalized learning: A guide for engaging students with technology*. International Society for Technology in Education.

[19] Kara M Gregory, An Sook Kim, and Alice Whiren. 2003. The effect of verbal scaffolding on the complexity of preschool children's block constructions. *Play and Culture Studies* 5 (2003), 117–134.

[20] Muhammad Usman Hadi, R Qureshi, A Shah, M Irfan, A Zafar, MB Shaikh, N Akhtar, J Wu, and S Mirjalili. 2023. A Survey on Large Language Models: Applications, Challenges, Limitations, and Practical Usage. *TechRxiv* (2023).

[21] Jamie J Jirout and Nora S Newcombe. 2015. Building blocks for developing spatial skills: Evidence from a large, representative US sample. *Psychological science* 26, 3 (2015), 302–310.

[22] Marjaana Kangas. 2010. Creative and playful learning: Learning through game co-creation and games in a playful learning environment. *Thinking skills and Creativity* 5, 1 (2010), 1–15.

[23] Enkelejda Kasneci, Kathrin Seßler, Stefan Küchemann, Maria Bannert, Daryna Dementieva, Frank Fischer, Urs Gasser, Georg Groh, Stephan Günnemann, Eyke Hüllermeier, et al. 2023. ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and individual differences* 103 (2023), 102274.

[24] Hlologelo Climant Khoza and Audrey Msimanga. 2022. Understanding the nature of questioning and teacher talk moves in interactive classrooms: A case of three South African teachers. *Research in Science Education* 52, 6 (2022), 1717–1734.

[25] Torkil Kollsker and Thomas JR Stidsen. 2021. Optimisation and static equilibrium of three-dimensional lego constructions. In *Operations Research Forum*, Vol. 2. Springer, 1–52.

[26] Igor Labutov, Sumit Basu, and Lucy Vanderwende. 2015. Deep questions without deep understanding. In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*. 889–898.

[27] Jungeun Lee, Suwon Yoon, Kyoosik Lee, Eunae Jeong, Jae-Eun Cho, Wonjeong Park, Dongsun Yim, and Inseok Hwang. 2024. Open Sesame? Open Salami! Personalizing Vocabulary Assessment-Intervention for Children via Pervasive Profiling and Bespoke Storybook Generation. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 1–32.

[28] Hilary E Miller and Vanessa R Simmering. 2018. Children's attention to task-relevant information accounts for relations between language and spatial cognition. *Journal of experimental child psychology* 172 (2018), 107–129.

[29] Shannon M Pruden, Susan C Levine, and Janellen Huttenlocher. 2011. Children's spatial thinking: Does talk about the spatial world matter? *Developmental science* 14, 6 (2011), 1417–1430.

[30] Geetha B Ramani, Erica Zippert, Shane Schweitzer, and Sophia Pan. 2014. Preschool children's joint block building during a guided play activity. *Journal of Applied Developmental Psychology* 35, 4 (2014), 326–336.

[31] Janet C Read, Stuart MacFarlane, and Chris Casey. 2002. Endurability, engagement and expectations: Measuring children's fun. In *Interaction design and children*, Vol. 2. Citeseer, 1–23.

[32] Siamak Shakeri, Cicero dos Santos, Henghui Zhu, Patrick Ng, Feng Nan, Zhiguo Wang, Ramesh Nallapati, and Bing Xiang. 2020. End-to-end synthetic data generation for domain adaptation of question answering systems. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. 5445–5460.

[33] Kayleigh Skene, Christine M O'Farrelly, Elizabeth M Byrne, Natalie Kirby, Eloise C Stevens, and Paul G Ramchandani. 2022. Can guidance during play enhance children's learning and development in educational contexts? A systematic review and meta-analysis. *Child Development* 93, 4 (2022), 1162–1180.

[34] Ivan Sysoev, James H Gray, Susan Fine, Sneha Priscilla Makini, and Deb Roy. 2022. Child-driven, machine-guided: Automatic scaffolding of constructionist-inspired early literacy play. *Computers & Education* 182 (2022), 104434.

[35] Lisa E Szechter and Lynn S Liben. 2004. Parental guidance in preschoolers' understanding of spatial-graphic representations. *Child Development* 75, 3 (2004), 869–885.

[36] Rachel R Van Der Stuyf. 2002. Scaffolding as a teaching strategy. *Adolescent learning and development* 52, 3 (2002), 5–18.

[37] Brian N Verdine, Roberta M Golinkoff, Kathryn Hirsh-Pasek, Nora S Newcombe, Andrew T Filipowicz, and Alicia Chang. 2014. Deconstructing building blocks: Preschoolers' spatial assembly performance relates to early mathematical skills. *Child development* 85, 3 (2014), 1062–1076.

[38] Deena Skolnick Weisberg, Kathy Hirsh-Pasek, and Roberta Michnick Golinkoff. 2013. Guided play: Where curricular goals meet a playful pedagogy. *Mind, Brain, and Education* 7, 2 (2013), 104–112.

[39] Dandan Wu, Hui Li, and Sheila Degotardi. 2021. Early spatial language development and education: a scoping review. *SN Social Sciences* 2, 2 (2021), 20.

[40] Wenjie Xu, Jiayi Ma, Jiayu Yao, Weijia Lin, Chao Zhang, Xuanhe Xia, Nan Zhuang, Shitong Weng, Xiaoqian Xie, Shuyue Feng, et al. 2023. Mathkingdom: Teaching children mathematical language through speaking at home via a voice-guided game. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–14.

[41] Ying Xu. 2020. Using conversational agents to foster young children's science learning from screen media. In *Proceedings of the 2020 ACM Interaction Design and Children Conference: Extended Abstracts*. 14–19.

[42] Ying Xu, Stacy Branham, Xinwei Deng, Penelope Collins, and Mark Warschauer. 2021. Are current voice interfaces designed to support children' s language development?. In *Proceedings of the 2021 CHI conference on human factors in computing systems*. 1–12.

[43] Yue Yu, Patrick Shafto, Elizabeth Bonawitz, Scott C-H Yang, Roberta M Golinkoff, Kathleen H Corriveau, Kathy Hirsh-Pasek, and Fei Xu. 2018. The theoretical and methodological opportunities afforded by guided play with young children. *Frontiers in psychology* 9 (2018), 1152.

[44] Xiajie Zhang, Cynthia Breazeal, and Hae Won Park. 2023. A Social Robot Reading Partner for Explorative Guidance. In *Proceedings of the 2023 ACM/IEEE International Conference on Human-Robot Interaction*. 341–349.

[45] Zheng Zhang, Ying Xu, Yanhao Wang, Bingsheng Yao, Daniel Ritchie, Tongshuang Wu, Mo Yu, Dakuo Wang, and Toby Jia-Jun Li. 2022. Storybuddy: A human-ai collaborative chatbot for parent-child interactive storytelling with flexible parental involvement. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–21.

Appendix

	Spatial Language Vocabularies
1. Spatial Dimensions	Big/Little, Long/Short, High/Low, Wide/Narrow, Thick/Thin, Skinny/Fat, Deep/Shallow, Full/Empty, Length, Height, Width, Depth, Volume, Capacity, Area.
2. Shapes	Circle, Square, Rectangle, Triangle, Oval, Semicircle, Polygon, Cube, Sphere, Cylinder, Cone.
3. Locations and Directions	At, To/From, On/Off, Ahead/Behind, Left/Right, In/Out, Between/Aside, Opposite, Position, Direction, Distance, Path.
4. Orientations and Transformations	Forward/Backward, Turn left/right, Upward/Downward, Flip, Rotate, Slide, Clockwise/-Counterclockwise.
5. Continuous Amount	Whole/Part, All/Half/Fraction, A lot/A little, More/Less, Same/Equal.
6. Deictics	Here, There, Where, This, That, Which.
7. Spatial Features and Properties	Line, Curve, Edge, Flat, Bent/d, Point, Acute angle, Obtuse angle, Right angle, Vertical, Parallel.
8. Pattern	Increase/Decrease, Before/After, Next/Last, First/Last, Order, Repeat, Pattern.

Table 4. Eight dimensions of spatial language and the corresponding vocabularies [7] used in the BrickSmart system.

	空间语言词汇
1. 空间维度	大/小, 长/短, 高/低, 宽/窄, 厚/薄, 粗/细, 深/浅, 满/空, 尺寸, 长度, 高度, 宽度, 深度, 体积, 容积, 面积
2. 形状	圆形, 正方形, 长方形, 三角形, 椭圆形, 半圆, 多边形, 立方体, 球体, 圆柱体, 圆锥体
3. 位置和方向	在, 从/到, 上/下, 前/后, 左/右, 里/外, 中间/旁边, 相对, 位置, 方向, 距离, 路径
4. 方向和变换	向前/后, 向左/右, 向上/下, 翻转, 旋转, 滑动, 顺时针/逆时针
5. 连续量	整体/部分, 全部/一半/几分之一, 多数/少数, 更多/少, 相等
6. 指示词	" 这里", " 那里", " 哪里", " 这个", " 那个", " 哪个"
7. 空间特征和属性	直线, 曲线, 边缘, 平面, 曲面, 点, 锐角, 钝角, 直角, 垂直, 平行
8. 模式	增加、减少, 之前/之后, 下一个/上一个, 第一个/最后一个, 顺序, 重复, 模式

Table 5. Eight dimensions of spatial language and the corresponding vocabularies [7] used in the BrickSmart system.

Received 20 February 20xx; revised 12 March 20xx; accepted 5 June 20xx

Steps	Goal	System Prompts
Step 1	Generate descriptions	Your task is to break down the story and scenes described by the child into several describable 3D objects, and output them sequentially into a structured string list object_list. For example, if the child says "A monkey with big eyes is climbing a tree," the output string list should be: object_list= Monkey, big eyes, action is climbing a tree, Tree. Please only break down and output based on the child's description without adding extra information or your own ideas. If the description lacks sufficient detail, output an empty string list object_list. You already have an initial list: {object_list}. Based on new descriptions or existing content, you need to refine and complete the entries in the list. Here is the conversation history so far: {chat_history}.
	Generate prompts for 3D models	You are an expert at crafting text prompts for generating 3D models, specializing in transforming children's imaginative words and narrated diaries into delightful, cartoon-style 3D models. Your prompts should focus on describing a single object rather than a scene, ensuring that the description is suitable for conversion into LEGO models. When rewriting user input, consider the following: The generated 3D model should avoid unnecessary details, should have a suitable center of gravity, and should clearly and independently represent a single entity. The appearance can be slightly enhanced to appeal to the aesthetic preferences of 6-8-year-old children. Ensure the model is suitable for construction using only basic LEGO bricks, emphasizing square and rectangular forms. Your output should be a detailed sentence or a list of descriptive words separated by commas, written in English.
Step 2	Understanding tutorial images	Based on the provided LEGO assembly tutorial image, thoroughly analyze and describe the current building task, outputting a string parameter instruction: (1) Describe the overall structure and design features of the LEGO model shown in the image. (2) Identify and describe the types of LEGO pieces and their colors appearing in the image. (3) Outline the assembly steps shown in the image, including any special assembly techniques or details that require special attention. (4) Use professional LEGO terminology to enhance the accuracy and professionalism of the description. (5) Ensure the text description accurately reflects the image content, with clear, professional, and detailed language to facilitate understanding of the assembly process.
	Vocabulary selection	You are a spatial language teacher, responsible for selecting spatial language tasks based on the task and student's proficiency. Use the following information: LEGO assembly tutorial description: {instruction}. Eight spatial language dimensions: ... User's current spatial language proficiency: {understand_level}, representing progress in each dimension (as a percentage). Select 3 suitable spatial language categories, focusing on those relevant to the LEGO task and where proficiency is lower. Output a list spatial_list with 3 elements, each from 0 to 7, representing category indices.
	Guidance generation	You are a family guide helping parents improve their child's spatial language. Based on the current LEGO tutorial, generate real-time prompts for parents to teach these words: Word: word_1, Stage: {stage_1}; Word: {word_2}, Stage: {stage_2}; Word: {word_3}, Stage: {stage_3}. The current tutorial is: {instruction}, showing a top view (current step) and a whole view (completed so far). Use the building task to teach the words, matching each to its learning stage. Example format: Word: Circle, Stage: Noun Explanation. Prompt: Explain that a circle has no corners, and all points on the edge are equally distant from the center. Example: Point out circular bricks or designs to help understanding.
Step 3	Guidance generation	You are an assistant helping parents improve their child's spatial language skills. They've built LEGO models: {objects}. Your task is to guide parents to move these models and have the child describe the actions to enhance spatial understanding. Output format: Vocabulary: The word to learn. Movement Example: Specific movement of the object. Parent Prompt: Example guidance for parents. Example: Vocabulary: Left/Right. Movement Example: Move the figure forward, then turn left. Parent Prompt: "Look, the figure turns left. Can you make it turn right?". For the num_words keywords: {keywords}, provide suggestions.

Table 6. The detailed prompts BrickSmart employs across three steps. Each step is tailored to enhance children's spatial reasoning and language skills through structured interactions and tasks.

Steps	Goal	System Prompts
Step 1	Generate descriptions	你的任务是孩子讲述的故事和场景分解成几个可描述的 3D 对象，并依次输出到结构化的字符串列表 object_list[] 中。例如，孩子说“大眼睛的猴子在爬树”，则输出字符串列表 object_list=[” 猴子，大眼睛，动作是正在爬树”, ” 树”]。请你仅按照孩子的描述来分解和输出，不要添加额外的信息或者自已的想法。如果孩子描述的内容不够详细，则输出空白字符串列表 object_list=[]。现在已经有了初步的列表信息：{object_list}。你需要根据新的描述或者已有内容，完善和补充列表中的条目。之前的对话历史如下：{chat_history}
	Generate prompts for 3D models	You are an expert at crafting text prompts for generating 3D models, specializing in transforming children’s imaginative words and narrated diaries into delightful, cartoon-style 3D models. Your prompts should focus on describing a single object rather than a scene, ensuring that the description is suitable for conversion into LEGO models. When rewriting user input, consider the following: The generated 3D model should avoid unnecessary details, should have a suitable center of gravity, and should clearly and independently represent a single entity. The appearance can be slightly enhanced to appeal to the aesthetic preferences of 6-8-year-old children. Ensure the model is suitable for construction using only basic LEGO bricks, emphasizing square and rectangular forms. Your output should be a detailed sentence or a list of descriptive words separated by commas, written in English.
Step 2	Understanding tutorial images	根据提供的乐高拼装教程图片，详细分析并描述当前积木搭建任务，输出字符串参数 instruction：(1) 图片展示的乐高模型的整体构造和设计特点。(2) 识别并描述图片中出现的乐高零件种类及其颜色。(3) 指出图片中的拼装步骤，包括任何特别的组装技巧或需要特别注意的细节。(4) 使用专业的乐高术语来增加描述的准确性和专业性。(5) 确保输出的文字描述准确反映图片内容，语言表达清晰、专业、详细，便于读者理解拼装过程。
	Vocabulary selection	你是一个空间语言教师，你的职责是根据任务和学生情况判断该学习的空间语言任务。请根据以下信息：乐高拼装教程描述 {instruction}. 空间语言的 8 个维度: ... 用户当前对空间语言的掌握程度是：{understand_level}，代表了 8 个维度分别的学习进度（以百分比表示）。判断最适合学习的 3 个空间语言类别，尽量挑选适合当前积木搭建任务且用户掌握程度不高的类别。输出长度为 3 的列表 spatial_list，每个元素的取值范围为 0-7，代表空间语言类别的索引。
	Guidance generation	你是一个家庭引导师，你的职责是帮助家长引导孩子并提升空间语言能力。你需要根据目前乐高搭建教程的步骤，实时为家长生成引导提示。当前需要学习的空间词汇和对应的阶段包括：1. 词汇：{word_1}，学习阶段：{stage_1}；2. 词汇：{word_2}，学习阶段：{stage_2}；3. 词汇：{word_3}，学习阶段：{stage_3}；当前乐高搭建的教程为：{instruction}，图片中包含俯视图（top view）-为当前步骤要搭建的积木块，和整体视图（whole view）-包含当前任务和之前已搭建好的所有。请理解当前的搭建任务，在引导搭建孩子的过程中学习以上三个词汇，分别符合对应的学习阶段，为家长生成提示和举例。示例格式输出：1. 词汇：圆形，学习阶段：名词解释提示：可以告诉孩子，圆形是一种没有角的形状，边上的每一点到中心点的距离都是一样的。示例：在搭建过程中可以找到圆形的积木块，或者积木块上的圆形图案，帮助他们理解。
Step 3	Guidance generation	你是一个帮助家长与孩子互动的助手，目标是提升孩子的空间语言表达能力。孩子和家长当前使用乐高积木搭建了几个乐高模型，包括：{objects}。你的任务是引导家长和孩子，通过让这些乐高模型“动起来”来描述它们的状态，从而增强孩子对空间概念的理解。家长可以移动搭建好的上述乐高模型，或乐高积木，并让孩子描述这些动作。请按照以下格式输出互动建议：词汇：要学习的空间语言词汇, 动态指令例子：给出移动物体的具体方法, 家长引导语示例：提供家长可以使用的引导语, 示例格式输出：1. 词汇：向左/向右; 动态指令例子：让小人向前走，然后向左转; 家长引导语示例：你看这个可爱的小人，他往前走，然后向左转啦。你能让小人向右转吗？针对下列 {num_words} 个关键词：{keywords}，逐个输出互动建议。

Table 7. The original Chinese version of prompts BrickSmart employs across three steps.

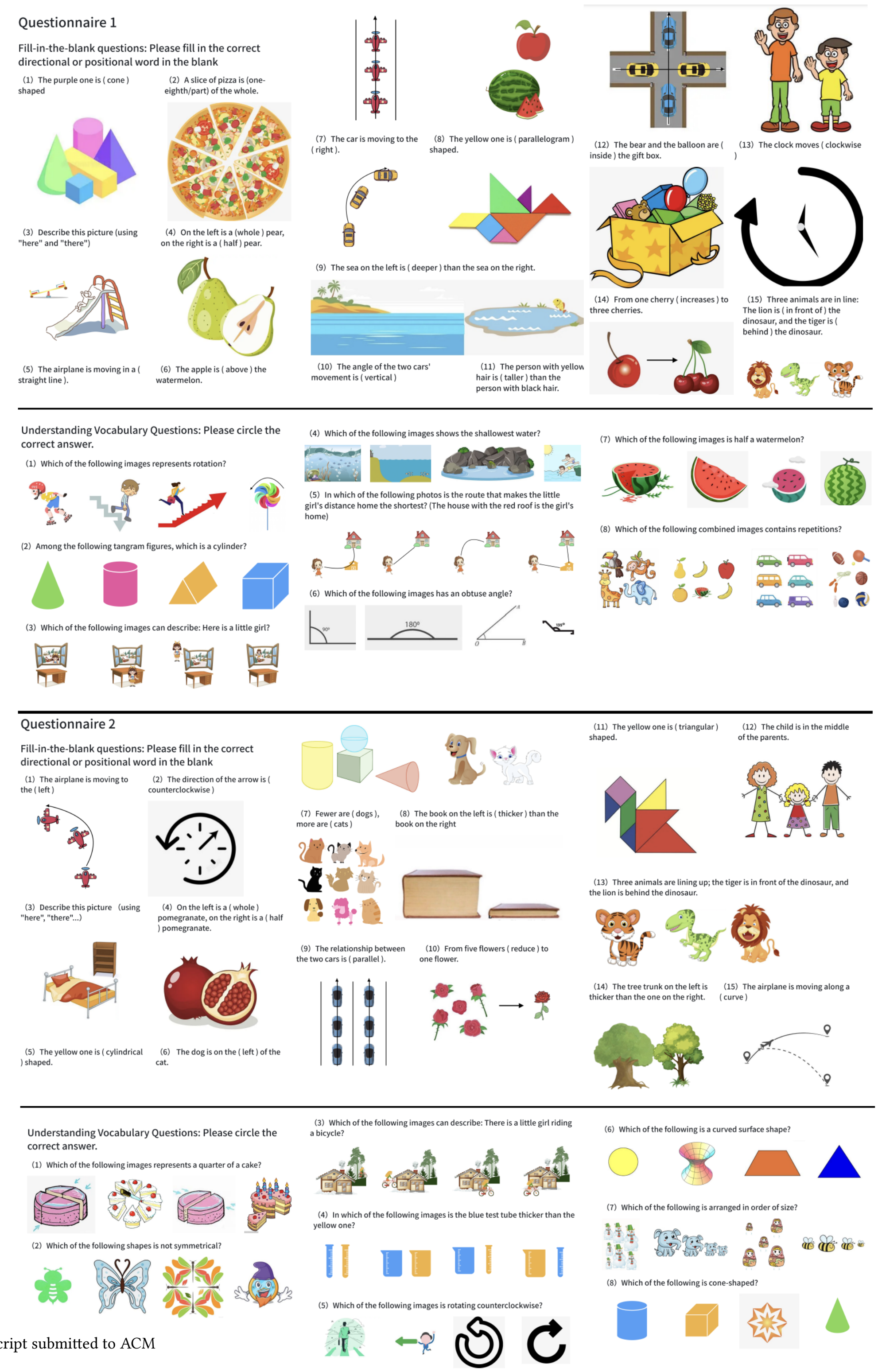

Fig. 10. Questionnaire of spatial language testing. The pre-test and post-test questionnaires are alternated between Questionnaire 1 and Questionnaire 2.