



# Designing LLM-simulated Immersive Spaces to Enhance Autistic Children's Social Affordances Understanding in Traffic Settings

Yancheng Cao

College of Design and Innovation,  
Tongji University  
Shanghai, China  
Institute for AI Industry Research,  
Tsinghua University  
Beijing, China  
yanchengc7@outlook.com

Menghan Chen

Institute for AI Industry Research,  
Tsinghua University  
Beijing, China  
chenmenghan0815@163.com

Min Liu

Institute for AI Industry Research,  
Tsinghua University  
Beijing, China  
2660991881@qq.com

Xin Tong

Thrust of Computational Media and  
Arts  
Hong Kong University of Science and  
Technology (Guangzhou)  
Guangzhou, China  
xint@hkust-gz.edu.cn

Yangyang He

Institute for AI Industry Research,  
Tsinghua University  
Beijing, China  
yangyanghe0821@gmail.com

Shanhe You

Institute for AI Industry Research,  
Tsinghua University  
Beijing, China  
u6692394@anu.edu.au

Chuan Luo

Tsinghua University  
Beijing, China  
chuanluobj@hotmail.com

Jing Liang\*

Tongji University  
Shanghai, China  
12046@tongji.edu.cn

Yonglin Chen

Institute for AI Industry Research,  
Tsinghua University  
Beijing, China  
yonglin0711@gmail.com

Yulin Qiu

Institute for AI Industry Research,  
Tsinghua University  
Beijing, China  
brantbaobao@outlook.com

Chen Zheng

Institute for AI Industry Research,  
Tsinghua University  
Beijing, China  
chen.zheng.psy@outlook.com

Jiangtao Gong\*

Tsinghua University  
Beijing, China  
gongjiangtao2@gmail.com

## Abstract

One of the key challenges faced by autistic children is understanding social affordances in complex environments, which further impacts their ability to respond appropriately to social signals. In traffic scenarios, this impairment can even lead to safety concerns. In this paper, we introduce an LLM-simulated immersive projection environment designed to improve this ability in autistic children while ensuring their safety. We first propose 17 design considerations across four major categories, derived from a comprehensive review of previous research. Next, we developed a system called AIRoad, which leverages LLMs to simulate drivers with varying social intents, expressed through explicit multimodal social signals. AIRoad helps autistic children bridge the gap in recognizing the intentions

behind behaviors and learning appropriate responses through various stimuli. A user study involving 14 participants demonstrated that this technology effectively engages autistic children and leads to significant improvements in their comprehension of social affordances in traffic scenarios. Additionally, parents reported high perceived usability of the system. These findings highlight the potential of combining LLM technology with immersive environments for the functional rehabilitation of autistic children in the future.

## CCS Concepts

- Human-centered computing → Empirical studies in HCI; Interactive systems and tools.

## Keywords

autistic children, LLM, immersive environment, social affordances, traffic settings

## ACM Reference Format:

Yancheng Cao, Yangyang He, Yonglin Chen, Menghan Chen, Shanhe You, Yulin Qiu, Min Liu, Chuan Luo, Chen Zheng, Xin Tong, Jing Liang, and Jiangtao Gong. 2025. Designing LLM-simulated Immersive Spaces to Enhance Autistic Children's Social Affordances Understanding in Traffic Settings. In

\*Mark corresponding authors.



This work is licensed under a Creative Commons Attribution 4.0 International License.  
*IUI '25, Cagliari, Italy*  
© 2025 Copyright held by the owner/author(s).  
ACM ISBN 979-8-4007-1306-4/25/03  
<https://doi.org/10.1145/3708359.3712142>

*30th International Conference on Intelligent User Interfaces (IUI '25), March 24–27, 2025, Cagliari, Italy.* ACM, New York, NY, USA, 19 pages. <https://doi.org/10.1145/3708359.3712142>

## 1 Introduction

Autism Spectrum Disorder (ASD) represents one of the most complex and least understood conditions in the field of neurodiversity [93], and it is also among its most prevalent forms [80]. Around one in every 36 children worldwide is diagnosed with ASD [85]. For these children, achieving social competence is challenging [63]. Previous research has revealed that autistic children face difficulties in reading facial expressions [26, 142], maintaining eye contact [43, 133], understanding gestures [5, 22], and discerning vocal tones [18, 118], among other areas. The social opportunities conveyed by these signals are referred to as *social affordances* [82]. The inappropriate understanding of *social affordances* by autistic children further diminishes their opportunities for social integration [6, 29] and consequently limit their chances to engage in and practice appropriate behaviors [83, 127], which are essential for functional rehabilitation [56, 146].

Current social skills training programs for autistic children face several challenges. Firstly, these interventions often require the presence of other participants, which increases the complexity of the training process [13]. This is particularly problematic given that autistic children frequently struggle to form friendships in mainstream schools [25, 113], while the time allocated for group training sessions in special education programs is typically limited [67]. Moreover, existing programs often focus on specific skills without adequately explaining the underlying reasons for social cues, raising concerns about the generalizability of these interventions [51, 107]. Additionally, the backgrounds and narratives in training scenarios are predefined, which restricts their capacity to capture the complexity of real-life social interactions and to achieve a degree of sophistication for the agent [14]. Compounding these issues is a notable lack of comprehensive synthesis regarding design considerations for interventions targeting autistic children [15].

In the area of supporting and training social skills for autistic children, the HCI community has made significant efforts, such as communication skill support [28, 97], social trigger spaces [147, 153], emotion recognition and expression [81, 131, 135], turn-taking and collaborative skills [11], among others [92]. In addition to support for these specific skills, previous research has also focused on particular contexts, such as the medical settings [52] and traffic scenarios [65]. Education for autistic children in crossing scenarios primarily employs VR technology [47, 76, 106, 116], along with touchable interfaces [125]. However, due to the complexity of social simulation, current immersive social affordance simulations for autistic children are still very limited, restricting the efficiency of training their social skills. Recently, the development of LLMs and their capabilities for social simulation present promising opportunities to address this issue [84, 90, 99, 101].

In this regard, this study aims to develop an LLM-simulated immersive space for autistic children that could facilitate social affordance understanding. In this paper, specifically, we hereby refer to **social affordance** as the range of **preferred behavior opportunities as determined by the combination of explicit and implicit social signals** [44, 156] and **general social norms** [109, 134]. In

certain specific contexts, such as traffic scenarios, where pedestrians are required to understand social signals and make decisions [105], this impairment can even pose risks to life and safety [37, 42]. To ensure the system is acceptable and effective, we conducted a comprehensive literature review to synthesize design considerations for educational interventions targeting autistic children. Then, we use a street-crossing scenario as an example, as the ability to read social signals and understand social affordances becomes crucial for safety when traffic lights are absent [117, 139]. Towards this end, we leveraged LLMs to empower this space. Through the LLM's social simulation capabilities [100, 101], we generate a rich variety of social intentions based on different driving styles, each embodying distinct social affordances. These are then expressed through social signals in the immersive space using multimodal representations. The system enables autistic children to interact safely with various car agents, allowing them to repeatedly practice, make mistakes, and receive corrective feedback.

Thus, the contributions of this paper are as follows:

- (1) Based on the comprehensive literature review, we identified four major categories of design considerations, encompassing a total of 17 of them regarding interactive educational technology for autistic children.
- (2) We developed an immersive space with LLM simulation to facilitate autistic children's understanding of social affordances, allowing multimodal perception and output.
- (3) The system can generate diverse driving intentions based on four different driving styles, with the underlying social affordances expressed through visual and auditory social signals.
- (4) A user experiment involving 14 autistic children was conducted to validate the system, demonstrating its usability and appeal to children, and confirming that it effectively helps them better choose the timing for crossing the street, which suggests an improved understanding of social affordances in traffic settings.

## 2 Related Work

### 2.1 Affordance and autistic children

Affordances refer to the opportunities for perception and action that the environment offers to animals [46]. The term, coined by Gibson, evolved when Donald Norman introduced it to design, distinguishing between perceived and actual affordances [95, 96] and highlighting potential everyday inconsistencies between them. Loveland et al. categorize environmental affordances into three levels: physical transactions, cultural preferences, and social signals reflecting others' meanings. Autistic children often struggle with the latter two types [82]. Difficulties with cultural affordances affect tool usage and may link to neurological issues [98]. Studies have explored interventions like electrical stimulation to enhance this understanding [79]. Understanding social signals involves grasping others' expectations, and research on helping autistic children in this area remains limited [109]. Perceptual issues with affordances negatively impact social participation of autistic children and contribute to adverse behaviors, reducing learning opportunities [55]. While emotional signal transmission is crucial for interpersonal coordination [54, 59, 119], autistic children's difficulties with social

affordances [57, 58, 60, 82] lead to challenges in social adaptation and independent living [24, 44, 66, 120].

## 2.2 Immersive environment for education

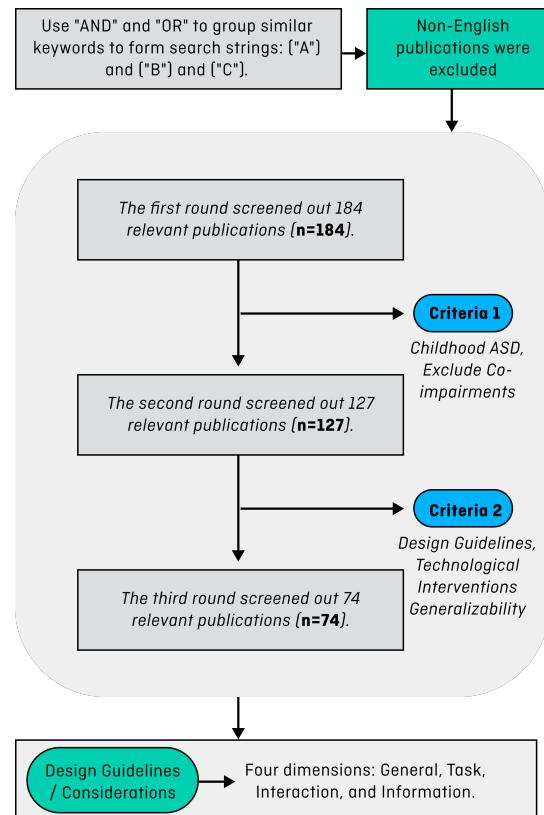
Autistic children often experience significant deficits in social understanding and skills. Immersive systems provide effective learning environments and support mechanisms to help address these challenges [32]. Studies have shown these technologies successfully improve social behaviors and enhance communication and emotional skills [12, 53, 81]. Through embodied interactions in virtual environments, children can safely explore new behavioral opportunities independently [53, 86]. These immersive settings provide a safe and inclusive space that reduces the hazards and unpredictability of real-life situations. Immersive environments are typically delivered through VR systems with Head Mounted Displays (HMDs) or projection-based systems like Cave Automatic Virtual Environments (CAVE) [15, 20]. While VR systems face limitations with autistic children often resisting headsets [78, 128], CAVE environments have proven effective for teaching safety skills, such as crossing streets or avoiding vehicles [136]. The current landscape of immersive environment systems presents significant research opportunities, particularly in integrating LLMs to enhance them with intelligent capabilities for memory, planning, and execution, though research in this area remains limited.

## 2.3 LLM-simulated social interaction

Large Language Models like ChatGPT have demonstrated remarkable ability in generating human-like responses [19]. LLMs excel in fundamental tasks like translation [130], conversation generation, and code writing, and have made significant advances in more complex domains such as autonomous decision-making and role-playing, as demonstrated by applications such as AutoGPT [151] and HuggingGPT [123] in task planning and execution. These advances have enabled practical applications across various domains - from creating character-aligned dialogues and simulating human behaviors in role-playing scenarios [101, 122, 141], to serving specialized functions like educational teaching assistance [27] and psychological counseling for individuals with high-functioning autism [33]. Recent research has developed innovative LLM applications. A project created a simulated job fair environment for training generative agents with enhanced communication capabilities [75], while another advanced social network simulation by modeling agents with emotional and interactive capabilities [45]. Previous research introduced an alignment learning approach that leverages simulated society interactions, providing collective ratings and iterative feedback [77]. These developments showcase LLMs' potential in mimicking human social interactions, suggesting a future where AI agents can participate in sophisticated social behaviors.

## 3 Design Considerations

Design considerations for autism-focused systems remain limited in current literature, particularly regarding interaction-focused system design and LLM-enabled immersive spaces. To establish comprehensive guidelines for LLM-enabled immersive systems for autistic children, we conducted a systematic literature review (Fig.



**Figure 1: The literature screening procedure in this study involved two rounds of evaluation. This process resulted in the retention of 74 articles as the basis for design considerations.**

- Given LLM's nascent status and limited literature on its application for autistic children, our review includes studies involving other emergent technologies for this demographic, particularly focusing on immersive technologies. Our keywords cover three dimensions: (1) target audience - autistic children; (2) application domains - education, intervention, and rehabilitation; (3) associated technologies - virtual reality and similar systems. We used Boolean operators to create search strings like: ("autistic children" OR "Autistic children" OR "Autistic Spectrum Disorder") AND ("Education" OR "Intervention" OR "Rehabilitation") AND ("Artificial Intelligence" OR "LLM" OR "Immersive environment" OR "Virtual reality"). We searched six databases: IEEE Xplore, ACM Digital Library, SpringerLink, Elsevier, ScienceDirect, and Google Scholar. Our team of five experts in engineering, psychology, and design conducted three screening rounds, reviewing titles, abstracts, and full texts. From 184 initial English articles, we applied two criteria: *Criterion 1* excluded studies of non-childhood ASD cases and those with co-occurring cognitive impairments, while *Criterion 2* identified articles addressing design considerations for technological interventions. The final 74 papers were analyzed to synthesize design guidelines across four dimensions: General, Task, Interaction, and Information.

### 3.1 General Guidelines

**3.1.1 Safe environment.** Autistic children often struggle to understand their surroundings, especially in terms of processing information, leading them to prefer environments that enhance their sense of control over interactions [68, 129]. Moreover, in the context of immersive environments, virtual reality devices such as head-mounted displays may induce negative experiences (e.g., dizziness or fatigue) for some autistic children [78, 128].

- General guideline 1: Provide an environment that makes children feel safe and comfortable.*

**3.1.2 Guidance and training.** It is essential to avoid lengthy and complex training sessions as autistic children often encounter difficulties in remembering and processing sequences [39, 50]. Given the attention comfort zone of autistic children, it is important to adjust the tone and pace of guidance appropriately and to be prepared to terminate training promptly if any discomfort arises [138]. Furthermore, studies on design considerations for autistic children emphasize the importance of providing instructions and the need for customized content adjustments tailored to the unique characteristics of this population [8, 38, 89].

- General guideline 2: Provide customized guidance and training in advance.*

### 3.2 Task Guidelines

**3.2.1 Objectives.** To facilitate participation for autistic children in current systems, it is crucial to establish a clear and explicit goal during task execution [9, 78, 143]. Bartoli et al. mentioned that the task objectives within the system should focus on a single task and a series of clear, manageable actions (e.g., "swinging arms to hit an object"), to aid cognitive processes related to organizing actions, thereby achieving the set goals [9].

- Task guideline 1: Provide clear and understandable task objectives.*

**3.2.2 Repeatability and predictability.** Unpredictability may elicit a host of adverse reactions, such as anxiety, in autistic children, necessitating the use of repeatable tasks to facilitate learning and maintain stability [9, 78]. System design should ensure that game tasks for autistic children are both repeatable and easily transitioned to higher levels to prevent discouragement. Whyte et al. emphasized the significance of repeatability and predictability, suggesting that repeating the same task can not only enhance mastery in autistic children but also provides the predictability they need, along with the anticipation of future behaviors [143].

- Task guideline 2: Employ repeatable and predictable tasks, allowing children to attempt and practice repeatedly.*

**3.2.3 Storytelling.** The narrative describes a technique used in the design of systems for autistic children, where some research has utilized this technique alongside multimedia to encourage autistic children to complete tasks and create their own narratives [30, 61, 72]. Certain story scenarios and cues can help improve the communication skills and symbolic functions of autistic children, thereby fostering their autonomy [152]. Soltiyeva et al. underscored the significance of storylines, noting that if autistic children become frightened by their surroundings, it can disrupt subsequent activities [128]. Hence, in spatial experiences, autistic children need to be guided and encouraged by stories from the outset. Moreover,

since autistic children often struggle to transition between activities, the introduction of stories can assist them in understanding this transition.

- Task guideline 3: Incorporate storylines to contextualize learning and enhance children's motivation.*

**3.2.4 Adjustable difficulty levels.** Given that each child exhibits distinct strengths and skill deficits, system design prioritizes addressing each child's individual abilities and needs. This advocating for personalized adjustments for different children, requiring systems that support a high level of personalization [21, 78]. System design should dynamically adjust the difficulty levels in alignment with the user's progression, and the intricacy of game tasks should be progressively increased [23, 78]. Tasks should allow for the introduction of incrementally more challenging tasks as autistic children assimilate certain rules within the system. This approach concurrently facilitates the enhancement of both motor and cognitive complexities [9, 34, 74, 143]. As autistic children acquire and reinforce appropriate skills, they typically manifest an evolving spectrum of needs, such as escalating requirements for motor, cognitive, and social capabilities [9, 74]. Furthermore, research in virtual environments has demonstrated that a gradual increment in task difficulty can yield superior outcomes for autistic children [69, 94, 129]. Pivotal Response Treatment (PRT) further suggests that the emphasis of each task should adapt to individual progress, accommodating more advanced objectives and needs [70, 91].

- Task guideline 4: Implement adaptable game difficulty levels, judiciously escalating complexity in correspondence with the personal capabilities of each individual child.*

### 3.3 Interaction Guidelines

**3.3.1 Multimodal input.** Multimodal systems that integrate interaction through voice, writing, touch, and other means offer distinct advantages over unimodal systems by leveraging redundancy and complementarity. This facilitates autistic children in conveying their interactive inputs to the system across multiple dimensions [140]. Previous research has also demonstrated the effectiveness of multimodal interventions in enhancing the learning and social communication skills of autistic children [10, 16].

- Interaction Guideline 1: Utilize multimodal perception at the input stage to collect behavioral data from the children.*

**3.3.2 Real-time feedback.** In interfaces designed for autistic children, feedback and signals related to the structure of interactive elements serve a crucial role in reinforcing the expected tasks for these children [39, 71]. Real-time feedback, in particular, aligns closely with their interaction expectations. Previous research has observed that positive real-time feedback can significantly motivate participants and influence their engagement in subsequent interactions [71, 73, 102].

- Interaction Guideline 2: Provide real-time feedback for each behavior exhibited by the children.*

**3.3.3 Reward mechanisms.** In the domain of system design for autistic children, providing positive reinforcement along with rewards is a strategy known as reward-based intervention [78, 143]. Bartoli et al. noted that offering reward stimuli to autistic children

upon the completing tasks that meet expectations can enhance motivation and engagement and may implicitly augment their skills [9]. Research indicates that some autistic children, particularly those with moderate to low functioning, may find intrinsically interesting rewards (such as videos or audio effects like entertaining animations or joyful music) more engaging than quantitative performance outcomes (e.g., scores) [9, 137]. Incorporating reward reinforcement in serious games designed for autistic children is an effective strategy that can enhance a variety of behaviors and skills [9, 40, 145].

- *Interaction Guideline 3: Provide positive reinforcement integrated with reward mechanisms.*

**3.3.4 Ease of input.** In system designs tailored for autistic children, simplicity and intuitiveness in control mechanisms are considered as superior design principles [15]. Complex and cumbersome input methods have been identified as sources of frustration for these children [39, 69, 73, 129]. Redundant controls can cause confusion among autistic children, who may be more prone to abandonment and disengagement than typical users [39, 144].

- *Interaction Guideline 4: Leverage simple and intuitive input methods to facilitate children's ease of use.*

**3.3.5 Devices.** For autistic individuals who may experience motor difficulties, it is advisable to employ a limited number of input devices [15, 39, 88]. Previous research has suggested that reducing the number of control buttons within input devices can benefit autistic children [39, 69, 129]. Moreover, autistic children may experience panic and resistance when using VR devices, with reports indicating that autistic children with anxiety disorders often fear and refuse to wear VR headsets [128]. Additionally, unfriendly methods of device usage can create barriers that hinder the normal progression of research activities.

- *Interaction Guideline 5: Utilize fewer devices that are also child-friendly to facilitate user input.*

## 3.4 Information Guidelines

**3.4.1 Vibrant colors and consistent style.** Autistic Children often process and remember visual information better than verbal information and are more sensitive to visual sensory stimuli than typically developing children [2, 34, 49, 108]. Research has suggested that systems designed for autistic children should incorporate more vibrant colors in the visual design to cater to this strength [3, 23, 126, 150, 154]. Moreover, a consistent color style can provide comfort and help prevent overstimulation.

- *Information Guideline 1: Employ vibrant colors and a consistent style.*

**3.4.2 Simplified information.** When designing graphics, sound, and text for systems aimed at autistic children, it is advised to avoid complex elements and lengthy compositions, as these may lead to distractions and, in severe cases, sensory overload [15, 34, 41, 103, 104]. Bartoli et al. noted that excessive visual stimuli can cause anxiety in autistic children because they may struggle to differentiate and interpret individual elements within a group, while too many auditory stimuli can create additional stress [9]. Moreover, research has highlighted that simplified graphics can enhance information processing for autistic children [41, 103, 104].

- *Information Guideline 2: Implement simplified graphics, sound, and text.*

**3.4.3 Cartoon characters.** Some studies have indicated that autistic children are more visually attentive to and engage in eye contact with cute and stylized cartoon characters than with images of real people [21, 111, 112, 157]. Furthermore, the use of simple 2D cartoon-like avatars in a system has been suggested as a more widely acceptable approach by autistic children compared to realistic human representations or complex characters [62, 78, 155].

- *Information Guideline 3: Utilize cartoon characters as virtual avatars instead of realistic representations.*

**3.4.4 Dynamic stimuli.** Prolonged static scenes in a system can result in a loss of attention in children, and in the case of autistic children, such static scenarios can even trigger motor rigidity [78, 143]. To counteract this, it is important to provide appropriate dynamic stimulation throughout the interaction to prevent the emergence of repetitive behaviors or motor rigidity in autistic children [9].

- *Information Guideline 4: Provide appropriate dynamic stimuli in the scenes to capture children's attention.*

**3.4.5 Multimodal prompts.** Autistic children often have difficulty processing external stimuli, which can affect their awareness of how their actions impact others and the environment. To address this, they may require multiple forms of prompts to guide subsequent actions [31, 78]. Previous studies have utilized a combination of media in the design of systems for autistic children to encourage interaction and promote autonomy [30, 61, 72]. Additionally, some research has integrated text, auditory, and visual cues to convey game instructions, with visual aids designed based on expert advice to assist autistic children to understand verbal commands [2, 78, 143].

- *Information Guideline 5: Provide prompts for children using various forms of information presentation, including text, sound, and visual cues.*

**3.4.6 Avoid overly sudden stimuli.** Some autistic children exhibit atypical fears of loud noises and fast-moving objects, leading to recommendations for excluding such elements from system designs intended for them [9, 61, 128]. Moreover, several studies have advised against incorporating sudden noises and unexpected, abrupt visual changes in the systems or educational content used by autistic children [39, 50, 138].

- *Information Guideline 6: Avoid the occurrence of overly unexpected stimuli in the scenes.*

## 4 System Design of AIRoad

In traffic scenarios without explicit traffic lights, understanding social affordances becomes crucial [82, 109]. To address this need, we developed AIRoad, a system specifically designed for street-crossing situations. This section first presents AIRoad's conceptual design and demonstrates its application through concrete examples. We then show how these implementations align with our previously established design guidelines, followed by a detailed discussion of the system's software and hardware components.

**Table 1: Design Guidelines for designing immersive interactive systems for autistic children.**

General Guidelines	
Provide an environment that makes children feel safe and in control.	[39, 50, 68, 78, 128, 129]
Provide customized guidance and training.	[8, 38, 39, 50, 89, 138]
Task Guidelines	
Provide clear and understandable task objectives.	[9, 78, 143]
Employ repeatable and predictable tasks, allowing children to attempt and practice repeatedly.	[9, 78, 143]
Incorporate storylines to contextualize learning and enhance children's motivation.	[30, 61, 72, 128, 152]
Implement adaptable game difficulty levels, judiciously escalating complexity in correspondence with the personal capabilities of each individual child.	[9, 21, 23, 34, 69, 70, 74, 78, 91, 94, 129, 143]
Interaction Guidelines	
Utilize multimodal perception at the input stage to collect behavioral data from the children.	[10, 16, 140]
Provide real-time feedback for each behavior exhibited by the children.	[39, 71, 73, 102]
Provide positive reinforcement integrated with reward mechanisms.	[9, 78, 137, 143]
Leverage simple and intuitive input methods to facilitate children's ease of use.	[15, 39, 69, 73, 129, 144]
Utilize fewer devices that are also child-friendly to facilitate user input.	[15, 39, 69, 88, 129]
Information Guidelines	
Employ vibrant colors and a consistent style.	[2, 34, 49, 108]
Implement simplified graphics, sound, and text.	[9, 15, 34, 41, 103, 104]
Utilize cartoon characters as virtual avatars instead of realistic representations.	[21, 62, 78, 111, 112, 155, 157]
Provide appropriate dynamic stimuli in the scenes to capture children's attention.	[9, 78, 143]
Provide prompts for children using various forms of information presentation, including text, sound, and visual cues.	[2, 30, 31, 61, 72, 78, 143]
Avoid the occurrence of overly unexpected stimuli in the scenes.	[9, 39, 50, 61, 128, 138]

## 4.1 Concept Design of AIRoad: Bridging Behavior and Intentions to Enhance Understanding of Social Affordance

Whereas neurotypical children naturally acquire social signals through daily exposure [1], autistic children often require structured practice [56]. We developed a virtual street-crossing scenario enabling safe, controlled practice [121], as illustrated in Fig. 2. AIRoad presents vehicle behaviors and their intentions through multiple modalities, including audio and visual cues, helping autistic children bridge behavior-intention connections in assessing driver yielding behavior. The system creates an LLM-simulated environment featuring street-crossing scenarios without traffic lights, displaying vehicles with various driving intentions through multi-faceted projections. Our driving intention generation is based on four styles from Taubman-Ben-Ari et al. [132]: dissociative, anxious, risky, and patient driving. Vehicle intentions are expressed through behavioral cues (speed and gestures) indicating yielding decisions, with LLM-generated narratives conveyed through speech synthesis. The LLM selects vehicle animations and generates corresponding intentions using literature-based prompts. Participants observe visual cues while hearing drivers' intentions through speech synthesis. During training, participants collect stars across road locations, requiring continuous interpretation of social affordances through multimodal information. This immersive space provides autistic children with a controlled learning environment [121]. Learning begins with enhanced scaffolding, while the system monitors performance to adjust difficulty and support levels. The LLM-powered adaptive space optimizes difficulty levels and provides multimodal scaffolding to facilitate skill development, ultimately enabling autistic children to comprehend social affordances effectively.

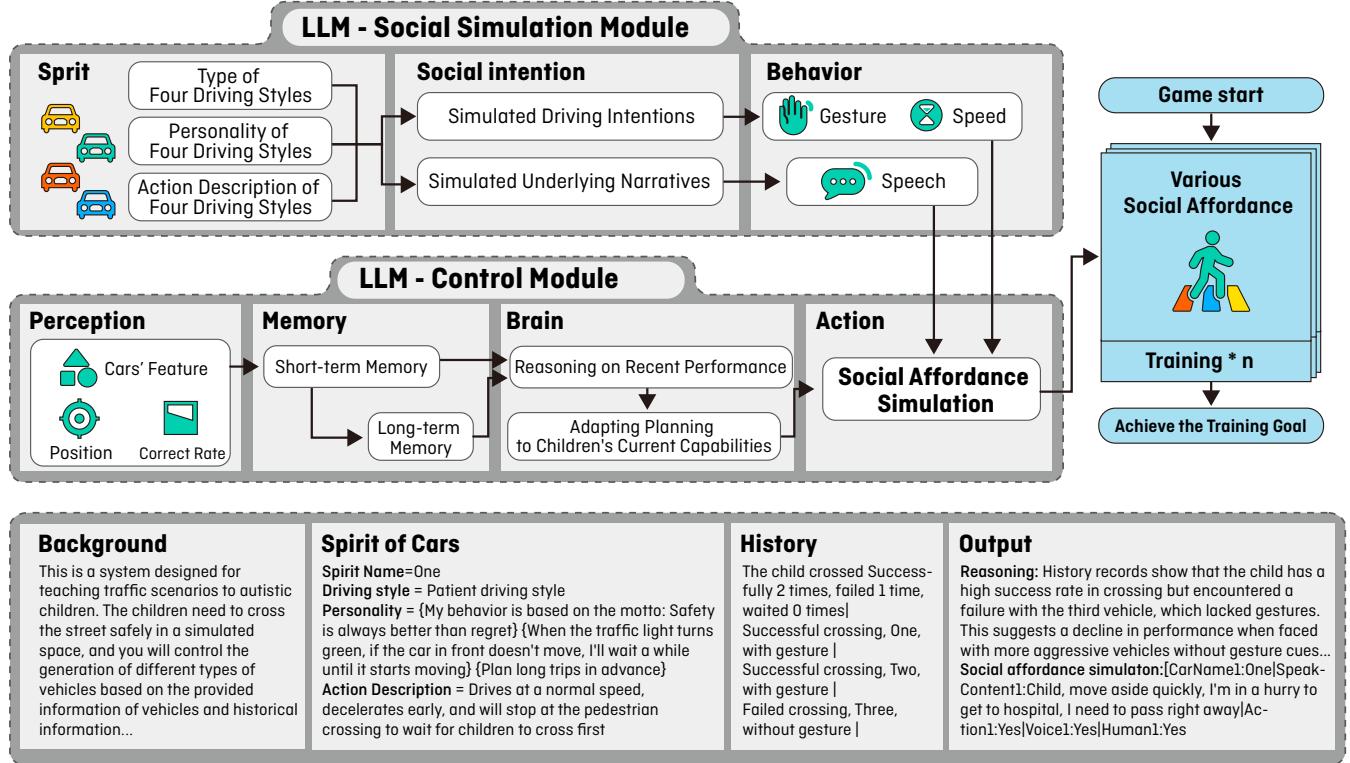
## 4.2 Example Scenarios to Illustrate Specific Functions

The system's core functions are illustrated through two contrasting examples. In the patient driving style, characterized by "Better safe than sorry," the vehicle slows down well in advance, and when participants show high error rates, the system displays a gesture encouraging pedestrians to cross. The LLM generates a matching narrative: "I'm heading to the supermarket to pick up a few things; there's no rush," which is conveyed through audio. Conversely, in the dissociative driving style, where the driver is "lost in thoughts or distracted," the vehicle maintains speed through the crossing without yielding. For participants with low error rates, no additional gesture cues are provided, while the LLM narrates: "I was on the phone and didn't see anyone on the road." These examples demonstrate how the system combines visual behaviors, adaptive gesture cues, and verbal intentions to help participants recognize different social affordances in crossing scenarios.

## 4.3 System Development Under Design Guidelines

The system development followed the aforementioned design considerations, denoted as DG1-DG17 based on their order in the table.

**4.3.1 System Development Under General Guidelines.** To meet **DG1**, the system enables users to input a nickname, facilitating personalized addressing in the generated voice prompts. We selected a simple intersection common in everyday environments as the context for our study. To address **DG2**, an introductory page was developed to facilitate interactions, allowing children to engage in dialogue with various vehicles and trees, thereby familiarizing themselves with the spatial layout before their experience with AIRoad. During the experimental phase, the facilitator provided a comprehensive overview of the game tasks, which is elaborated upon in the subsequent section.



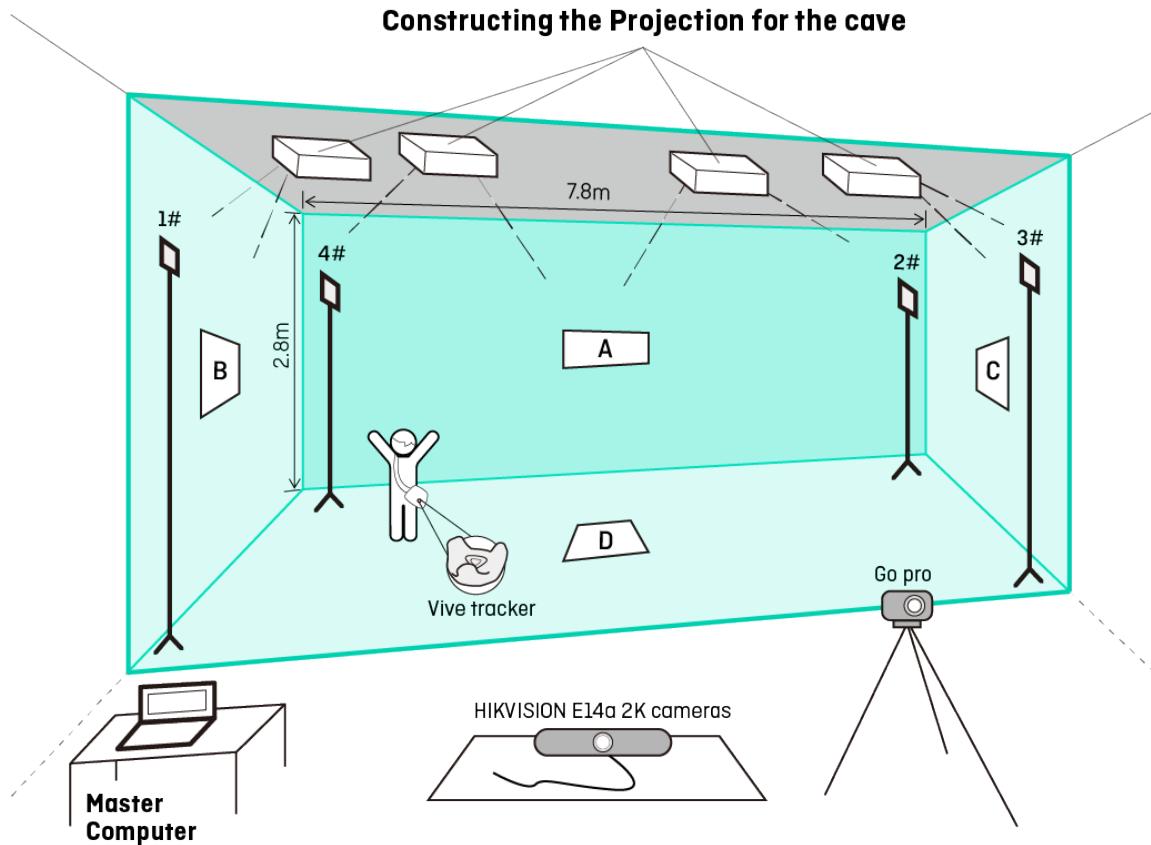
**Figure 2: System framework of AIRoad. The social affordance simulation conducted by the LLM is detailed in the social simulation module. The control of game rules by the LLM is outlined in the control module. Key prompts and outputs related to the LLM are also displayed below.**

**4.3.2 System Development Under Task Guidelines.** To align with **DG3**, we designed a clear and engaging objective: participants must collect stars while safely crossing the road, with emphasis on both speed and safety. In addressing **DG5**, we leveraged LLM's social simulation capabilities to generate diverse driving intention narratives, presenting them as engaging stories that enhance autistic children's engagement and understanding. Following **DG4**, the system implements repeated road-crossing scenarios, allowing participants to progressively improve their ability to interpret vehicle driving intentions while learning from their mistakes. To satisfy **DG6**, the system tracks participant errors through short-term and long-term memory, allowing the LLM to adjust difficulty levels dynamically. These adjustments include varying levels of cognitive support (such as gesture cues and voice-over narratives) and environmental challenges (like vehicle speeds and pedestrian interactions).

**4.3.3 System Development Under Interaction Guidelines.** To meet **DG7** and **DG11**, AIRoad utilizes a portable Vive tracker and a lightweight microphone as input devices, facilitating multimodal input while ensuring ease of transport. To address **DG10**, we minimized the operational tasks required of participants; the tracker functions without intervention, and the microphone can

be activated with a single button press. To fulfill **DG8**, all feedback animations are based on real-time position recognition. In accordance with **DG9**, the screen provides live updates on the number of stars collected and those yet to be obtained. Additionally, prior to the start of the experiment, participants were informed about the rewards system, which included sticker rewards distributed throughout the process.

**4.3.4 System Development Under Information Guidelines.** To meet **DG12** and **DG14**, the system's visual design incorporates a cartoonish style for vehicles and environments, utilizing vibrant colors that are appealing (as illustrated in Fig. 4). To address **DG13**, all graphics are designed to avoid excessive detail, and we have specified in the LLM prompts that generated speech should not exceed 25 words. To fulfill **DG15**, we created dynamic animations for moving vehicles, talking trees, collectible stars, and interfering pedestrians to stimulate engagement among autistic children. To meet **DG16**, the immersive space conveys prompts through multimodal elements such as vehicle animations, collision sounds, and narrative dialogues. Lastly, to address **DG17**, the vehicles encountered during training are positioned in fixed locations, and the voice prompts are generated using gentle tones to minimize sudden stimuli for autistic children.



**Figure 3: AIRoad is an AI-enabled immersive educational space tailored for autistic children. It facilitates autistic children’s learning about social affordances in complex traffic scenarios through the construction of virtual transportation settings.**

#### 4.4 System implementation

We developed a Unity-based virtual environment with LLM-controlled assets, featuring safety zones and dual crosswalks. The space is equipped with four VR stations that track users’ real-time positions to evaluate their behaviors during simulated vehicle crossing scenarios. This immersive environment captures user behaviors and transmits logs to LLMs, which then generate standardized text formats to orchestrate traffic simulations using predefined Unity assets. To enhance acceptance among autistic children, we chose projection as the primary medium, with detailed rationale provided in Table 1.

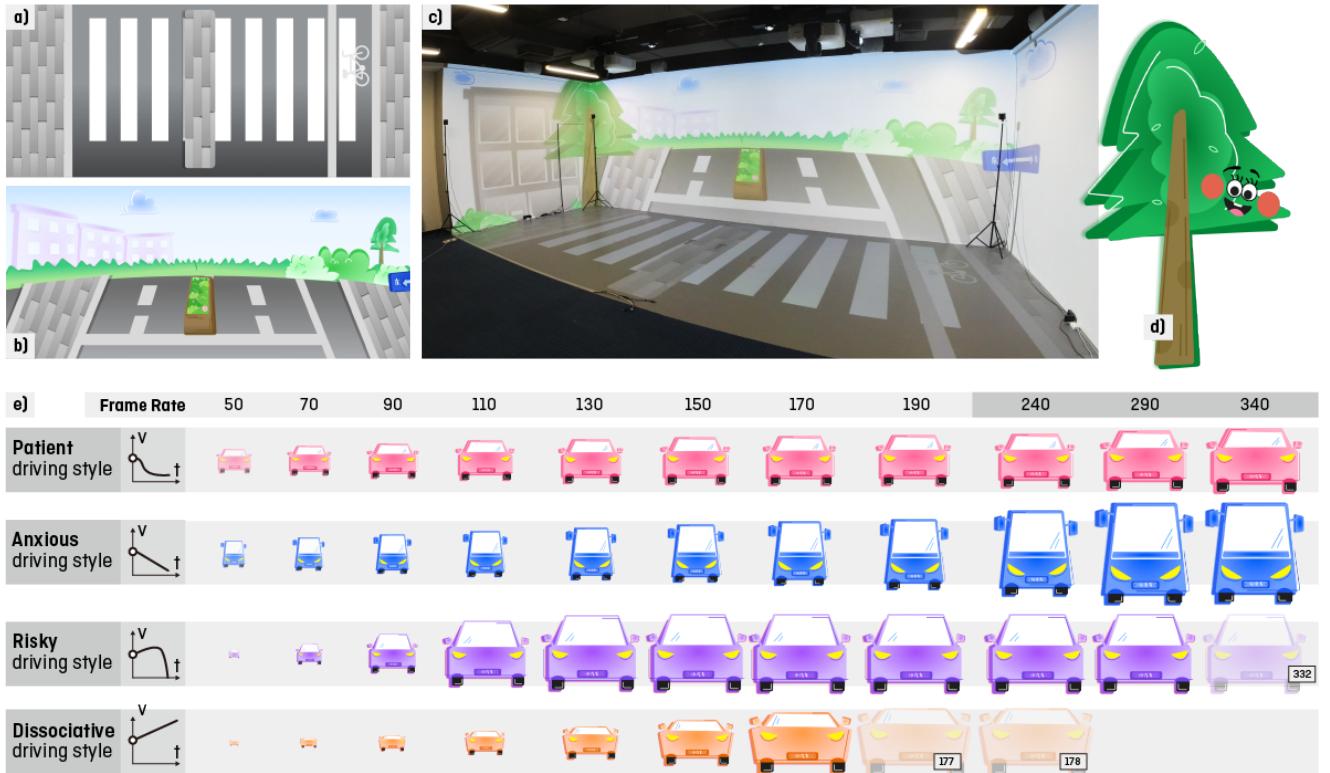
**4.4.1 Hardware and space setting.** As shown in Fig. 3, AIRoad creates an immersive experience using a four-projection system with HTC Vive base stations mounted at the four upper corners, and the HTC Vive trackers are used to track the player’s real-time position in virtual space. Notably, to avoid children participants feeling uncomfortable with wearing trackers, we strapped the tracker to a backpack and designed with a playful appearance, making them appealing and acceptable to most children so that they would be willing to carry a backpack while playing. Fig. 3 shows that the

projections displayed in faces A and D are of equal size (7.8 m in length and 2.8 m in width), while faces B and C are also of equal size (2.8 m in length and 2.8 m in width). Additionally, the experimenter uses a master computer to remotely control the computer connected to the projection system and has set up a *GoPro* and a *HIKVISION E14a 2K* camera to record the experimental process.

**4.4.2 Technical framework.** AIRoad utilizes a prompt-driven decision-making approach for LLMs, leveraging prompts to describe scenarios and requirements, thereby facilitating the decision-making and thought processes of the LLMs. The related source code and associated prompts are available at <https://github.com/Minadocyc/AIRoad.git>

**LLM Model Selection.** GPT-4 Turbo was selected as the underlying model after comparative testing with other models such as GPT-3.5 and ERNIE. At the time of development, it was OpenAI’s most capable model. It demonstrated superior performance in instruction compliance, social intention generation, and response time - particularly important given the lengthy prompts used.

**Prompt Development.** Our system interfaces with GPT-4 Turbo through OpenAI’s API, utilizing multiple locally stored prompt files for system configuration. These files include Background (system



**Figure 4:** a) Projection on the ground; b) Projection on the wall; c) Overall real-life overview; d) Other cartoon elements; e) Creation of animation effects and vehicle driving style (30 frames per second).

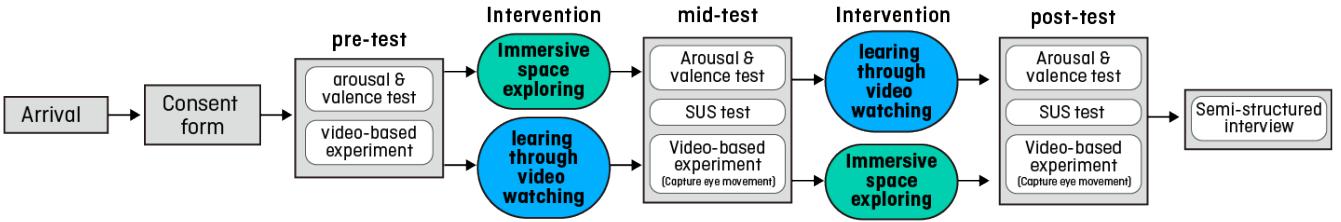
context); Tool (available tools and output format standards); Social (detailing characteristics of four vehicle types); and History (performance history and recent logs). During operation, these files are consolidated into a comprehensive prompt, sent to the LLM for decision analysis, which then generates simulated traffic scenarios.

**History Module.** The History Module stores log information that encapsulates the entire scenario context, also known as the context of scene changes. Changes in player positions and scene content trigger updates that are logged into the History Module. This comprehensive context, provided by the History Module, ensures that the LLMs have all the necessary information at their disposal when making decisions. When LLMs are called upon to make decisions, a concatenation of all logs and prompts serves as the input for the LLMs.

**Social Module.** A Social Module is an interactive entity capable of executing commands, producing animations, and generating unique sounds. These Module are defined through personified characteristics using JSON templates, with customizable attributes including type, personality, position, responsibilities, actions, and voice parameters. For example, we implemented a pink toy car character designed to be cute and lively to assist children in safe street crossing, with its voice synthesized using a trained VITS(Conditional Variational Autoencoder with Adversarial Learning) model to generate gentle female speech with distinctive tonal qualities.

**Real-time feedback.** Scene dynamics include participant positions, gesture information and the cars' social behavior. Due to LLMs' token limitations, the scene is divided into areas, with each area having a specific text description for functions. Player movements are updated in the memory model when position changes occur, and these updates are triggered either by player actions or scene changes. For decision-making, the LLM input combines memory model data with predefined prompts. The LLM will output the content that needs to be executed in a specific format that dictates Spirit interactions, animations, and vehicle generation based on participant engagement and appropriate difficulty levels. Finally, the output text is disassembled through the script and the corresponding animation or voice playback is executed.

**Synchronization of virtual and real space.** Within Unity, four cameras are set and orthographic mode is used to facilitate parallel projection. The views captured by these cameras are then mapped onto the projector's output using Spout, effectively translating the virtual scene into the physical space. Additionally, the positions of tracked participants are linked to a capsule object within Unity, allowing interaction between the players' locations and the colliders in the scene. To ensure accurate positional information about the players within virtual and real space, the vectors tracked by the Tracker are multiplied by  $n$  and adjusted by an offset  $b$ . These two values will be manually calibrated every time the system starts.



**Figure 5: The experimental procedure of the study**



**Figure 6: a) Children are participating in a video-based experiment; b) Completing a quiz on social affordance.**

## 5 User Experiment

After implementing the system, we conducted user experiments to evaluate its impact on real users. The primary research questions include the following three:

- RQ1 To what extent does AIRoad demonstrate usability for autistic children? How do these children perform within the training?
- RQ2 What are the engagement behaviors and emotional responses of autistic children when training with AIRoad?
- RQ3 Can AIRoad effectively enhance the understanding of social affordances in traffic settings among autistic children?

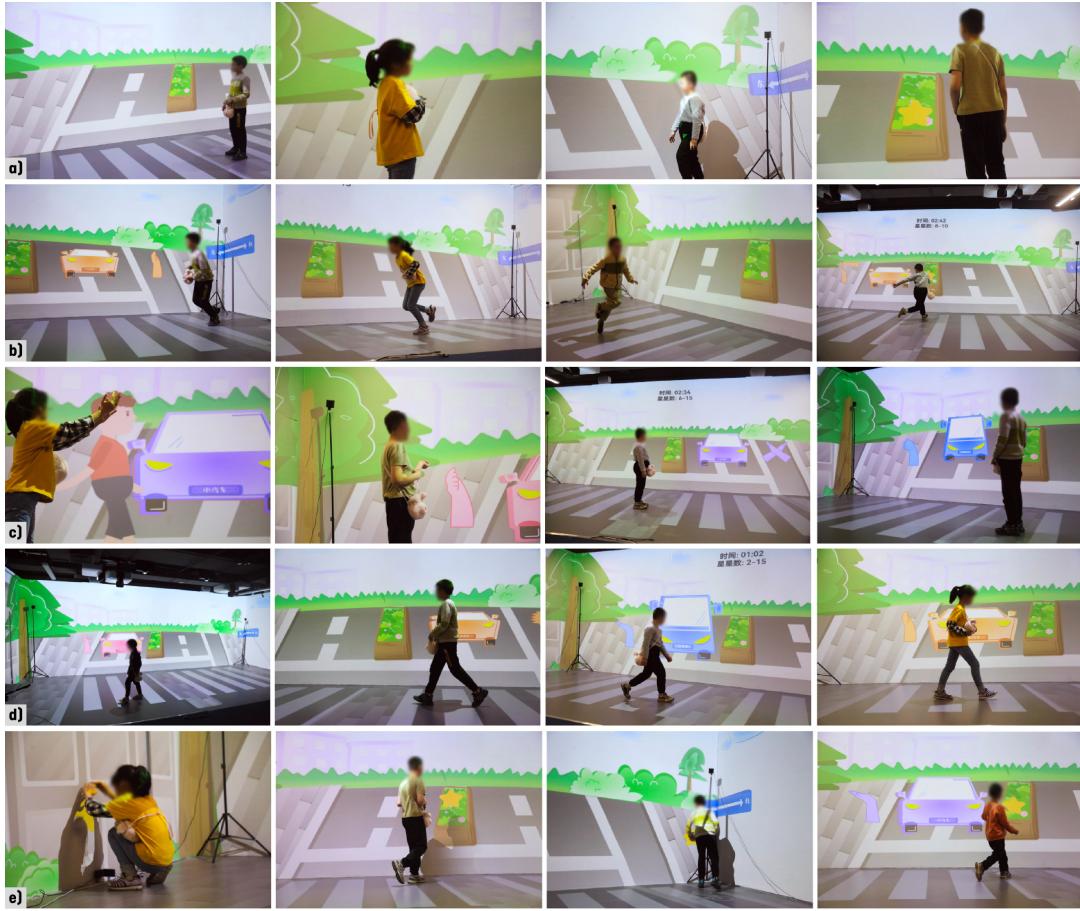
### 5.1 Participants and Ethical Approval

In our study, we recruited a total of 14 participants aged 6 to 12 years ( $M = 8.79$ ,  $SD = 2.01$ ), among whom 5 were recommended by a qualified special education institution, and the remaining 9 were recruited through a combination of online and offline methods. For the 9 participants, the recruitment process was structured and rigorous. Parents were required to provide self-reports on their children's conditions and complete the Autism Behavior Checklist (ABC). Additionally, parents voluntarily provided diagnostic reports from qualified institutions or hospitals, or professional opinions from certified specialists. This recruitment process enabled us to gather a suitable sample of participants while ensuring they met the necessary criteria for the study. This study was approved by the Ethics Committee of Tsinghua University.

### 5.2 Experimental Design and Experimental Procedure

We conducted a within-subject study comparing two experimental conditions: AIRoad training and video-based tutorials (Fig. 5). To counterbalance potential order effects, participants were randomly divided into two groups, experiencing either AIRoad or the video-based tutorial first. In the AIRoad condition, children first familiarized themselves with the virtual environment, then engaged in an immersive road scenario where they collected stars appearing in sequence (Fig. 7). The task was complete upon reaching a predetermined number of stars. The interface displayed progress and remaining time to help children track their status. In the video condition, participants watched curated road safety videos<sup>1</sup>. The experimental procedure consisted of three test sessions (pre-, mid-, and post-test), interspersed with two training processes, and concluded with a semi-structured interview. The procedure was as follows: (1) parent consent and briefing, (2) pre-test, (3) first training intervention (Fig. 6), (4) mid-test using the same materials as pre-test, (5) parent completion of System Usability Scale (SUS) [17], given children's limited compliance in answering such questions, (6) second training condition, (7) post-test using the same materials as mid-test, and (8) concluding interview. Throughout the study, each child was assisted by two experimenters, with parents nearby to address any special circumstances. All sessions were video recorded.

<sup>1</sup><https://www.youtube.com/watch?v=yEP3pws5lNQ&t=1s>



**Figure 7:** The procedure of immersive space exploring: a) Children observe their surroundings in AIRoad; b) Upon first seeing an approaching vehicle, they nervously run to the other side of the road; c) They observe and learn the different gestures and voice prompts of various vehicles; d) They gradually understand the intentions of different vehicles and begin to cross the road calmly; e) They successfully collect the stars and complete the mission.

### 5.3 Measurement

The study measurements consisted of three phases: test scores from pre-, mid-, and post-assessments, behavioral data logged by the AIRoad system, and post-hoc analysis of video recordings. This multi-faceted approach enabled thorough analysis of participants' learning progress and interactions.

- (1) **Video-based Experiments.** We recorded a set of videos depicting real street scenarios and asked autistic children to determine the appropriate moment to cross the street, aiming to evaluate participants' understanding of social affordance. Three authors collaboratively predetermined a safe period for each video, during which it was deemed safe to cross the street based on traffic light signals and passing vehicles. The accuracy and reaction time were computed according to these safe periods. A correct response was recorded if the child selected to cross the street within the predetermined safe period. The reaction time was measured as the duration between the start of the safe period and the participant's selection time.

(2) **Level of Arousal and Valence.** To assess emotional changes in children before and after the training condition, we employed the Valence Arousal Scale, derived from Russell's circumplex model of affect. This model posits that emotion is composed of two bipolar and orthogonal dimensions [114, 115]. Each scale ranged from -2 to 2, using integer values.

(3) **SUS Scales.** To measure system usability, we employed an adapted version of the System Usability Scale (SUS[17], see Appendix). Due to the limited compliance of autistic children in responding to such questionnaires, we asked parents to complete this scale on behalf of their children.

(4) **In-game Log Data.** To evaluate autistic children's performance in the AIRoad training, we recorded road-crossing accuracy based on log data. Accuracy was computed through real-time position detection; a correct result was registered if the child was waiting or had arrived at a safe region when the car-leaving animation initiated. An incorrect result was recorded when the child collided with a vehicle.

**Table 2: Manual coding scheme for Activity Participation and Emotional Activity**

Codes	Descriptions
Activity Participation	Engrossed: Follow instructions, participate in activities according to rules or have a visual gaze for major objects, sustained for at least 3 seconds counts as one instance
	Distracted: Ignore instructions, avert or shift eyes to extraneous objects, and unable to participate in activities according to rules
Emotional Activity	Emotional Expression: Exhibit associated emotions (pleasure, delight, agitation, excitement, nervousness, etc.)
	Problematic Behavior: Exhibit severe problem behavior (crying, emotional breakdowns, aggressive behavior, etc.)

(5) **Video Data and Analysis.** The entire experiment was recorded using a GoPro and a 2K camera (as shown in Fig. 3). These video recordings were primarily utilized for manual coding to analyze children's engagement and emotional expression [35]. The coding scheme was developed based on and incorporated elements from Gong et al. [48] and Wu et al. [148]. All children's recordings were coded. Two trained coders independently coded these videos, achieving an inter-rater Pearson correlation greater than 0.99.

## 5.4 Results

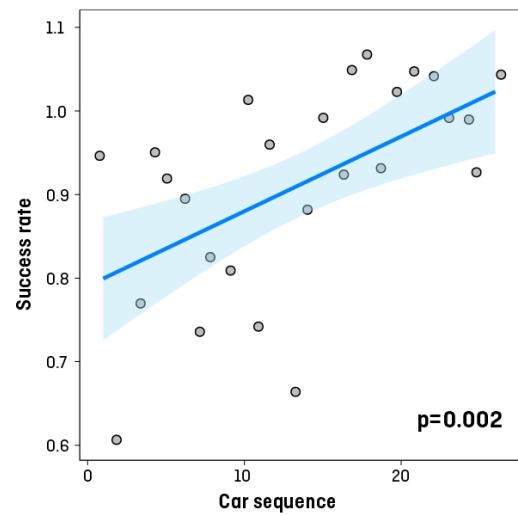
**5.4.1 RQ1: System usability and participants' performance.** The results of the *System Usability Scale (SUS)* are summarized in Table 3. The scores indicate that the usability of AIRoad was excellent ( $>85.0$ ), acceptable ( $>70.0$ ), and achieved an A+ level ( $>84.1$ ) according to Bangor et al.'s empirical evaluation [7]. As evident from the data presented in Table 3, AIRoad consistently scored higher than the video tutorial condition. Notably, AIRoad's score exceeded 85, placing it in the "Excellent" category. These findings address RQ1, suggesting that parents of autistic children perceive AIRoad as an acceptable and highly usable tool for this population.

**Table 3: System Usability Scale**

Experiments	AIRoad	Video-tutorial
Mean	85.71	80.36
SD	15.32	18.11

**In-game performance during experimental training.** We conducted a logistic regression analysis to evaluate the behaviors of children in AIRoad, particularly their ability to safely cross the street. As shown in Fig. 8, the analysis revealed a significant correlation between the training process and the rate of safe crossings ( $p = 0.002$ ). This result addresses RQ1, indicating that through repeated experiences in the game training, autistic children were able to gradually reduce their error rates and demonstrate improved performance. This finding may also suggest insights related to RQ3, indicating that autistic children may have gained a better understanding of social affordances. However, it is also possible that this improvement stems from their increasing familiarity with the game itself.

**5.4.2 RQ2: Engagement and Emotional Responses in AIRoad Training.** To examine the differences in children's engagement

**Figure 8: The relationship between the rate of participants successfully crossing the road and the number of cars used in training.**

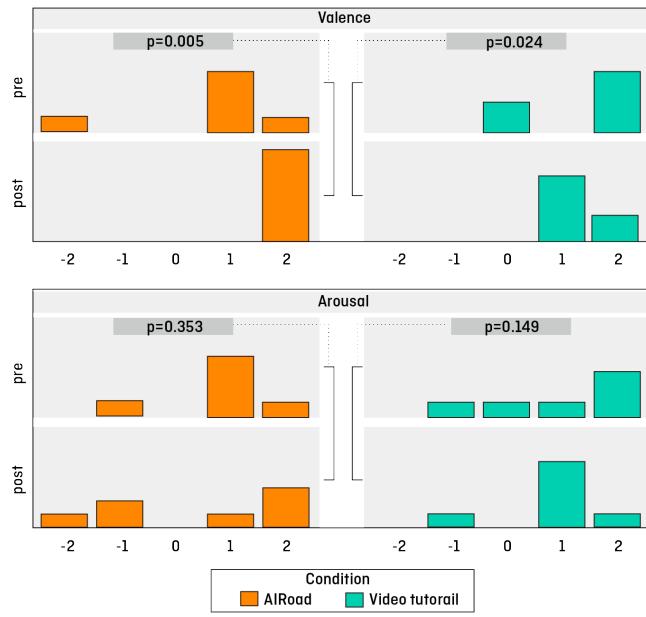
and emotional experiences between AIRoad and Video-tutorial, we employed a paired-sample T-test for video coding analysis. Analysis of the *Video coding* from the children's tests, as presented in Table 4, indicates that participants in the AIRoad condition demonstrated higher levels of activity participation and emotional engagement.

Regarding the Activity Participation Results, the mean of engrossing behavior in AIRoad condition was 55.615 ( $SD = 2.980$ ) and of engrossing in Video-tutorial condition was 42.000 ( $SD = 18.478$ ). This difference was statistically significant according to a paired-samples t-test ( $t(13) = 2.281, p < .05$ ). And the mean of distracting behavior in AIRoad condition was 4.385 ( $SD = 2.980$ ) and of distracting behavior in Video-tutorial condition was 18.000 ( $SD = 18.478$ ). This difference was statistically significant according to a paired-samples t-test ( $t(13) = -2.281, p < .05$ ).

In terms of Emotional Activity Results, the mean of emotional expressions in AIRoad condition was 1.539 ( $SD = 2.015$ ) and of emotional expressions in Video-tutorial condition was 0.227 ( $SD = 0.518$ ). This difference was statistically significant according to a paired-samples t-test ( $t(13) = 4.719, p < .001$ ). And the mean of problematic behavior in AIRoad condition was 5.500 ( $SD = 3.702$ ) and of problematic behavior in Video-tutorial condition was 2.546

**Table 4: Children's t-test results of video coding**

Data	Variable	AIRoad	Video-tutorial	Value	P
Activity Participation	Engrossing	M = 55.615 (SD = 2.980)	M = 42.000 (SD = 18.478)	t = 2.281	0.046*
	Distracting	M = 4.385 (SD = 2.980)	M = 18.000 (SD = 18.478)	t = -2.281	0.046*
Emotional Activity	Emotional Expression	M = 1.539 (SD = 2.015)	M = 0.227 (SD = 0.518)	t = 4.719	0.001**
	Problematic Behavior	M = 5.500 (SD = 3.702)	M = 2.546 (SD = 2.318)	t = -1.330	0.213

**Figure 9: The result of Valence and Arousal Scale.**

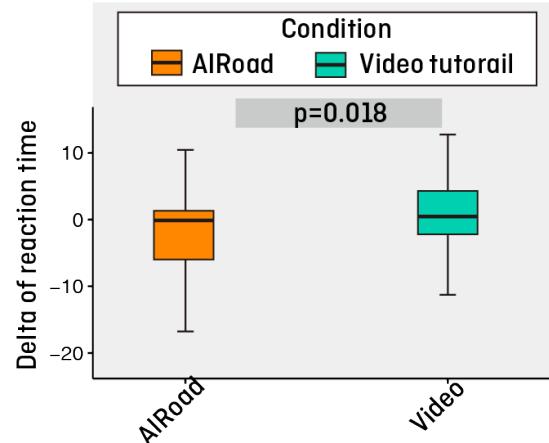
(SD = 2.318). This difference was not show statistically significant according to a paired-samples t-test ( $t(13) = -1.330$ ,  $p > .05$ ).

These findings suggest that children exhibit more positive affective states and increased attentional focus when engaged in AIRoad training than video tutorial. These results directly address RQ2, providing evidence that the AIRoad system we developed can offer autistic children a more engaged experience and promote positive emotional states.

We conducted a statistical analysis of the results from the **Valence and Arousal Scale**. As shown in Fig. 9, we conducted the Fisher test on the Valence and Arousal Scale. The results reveal a significant improvement in children's valence after undergoing AIRoad training ( $p < .01$ ). Conversely, following video-tutorial educational learning, a significant decrease in children's valence was observed ( $p < .05$ ). In contrast, no significant changes in arousal levels were detected in autistic children before and after the training with either educational technology ( $p = .353 / .149$ ).

This result addresses RQ2, indicating that AIRoad can bring positive emotions to autistic children, although there are no significant differences in the intensity of these emotions. Meanwhile, traditional video tutorials may lead to negative emotions in autistic children.

**5.4.3 RQ3: Effects on Enhancing the Understanding of Social Affordances. Video-based Experiments.** To assess the impact of training on autistic children's response times to video stimuli, we calculated the difference in reaction times before and after the training intervention. This measure is referred to as the *Delta of Reaction Time*. The result of *Delta of Reaction Time* can be seen in Fig. 10. The mean for the AIRoad condition was -2.20 (SD = 6.67), while for the video-tutorial condition it was 0.18 (SD = 6.00). The Wilcoxon rank-sum test indicated a statistically significant difference between the two groups ( $W = 2343$ ,  $p < .05$ ), suggesting that the *Delta of Reaction Time* in the video tutorial condition was significantly higher than in the AIRoad condition. These findings indicate that after AIRoad training, participants were able to respond more quickly to the appropriate time to cross the street. This also partially addresses RQ3, suggesting that autistic children exhibited a better understanding of social affordance in the context of the street-crossing scenario following their experience with AIRoad. Participant 3 noted, "But some cars, even though they're far away, they drive so fast and zoom; they're here in no time. So, just because they're far doesn't mean it's safe to go... And if they're close, it's not always okay; I gotta see if they're speeding fast or not." Participant 7 also mentioned, "If a car's going fast, it means they're in a hurry for something important, like going to work or to the hospital... And if a car's going slow, it means they don't have anything important going on." These qualitative results further demonstrate that autistic children gradually connected vehicle behavior with its underlying intentions through repeated practice in AIRoad training, enhancing their understanding of social affordance within traffic settings.

**Figure 10: Reaction Time**

## 6 Discussion

### 6.1 Large Language Models as Special Educational Support Tools for Autistic Children

LLMs have shown utility beyond common applications, as this study explores their potential in special education for autistic children. Leveraging autonomous decision-making and social simulation capabilities, these models offer promising solutions through consistent and adaptable learning tools, reducing burdens on educators. Their adaptability enables customized feedback and guidance based on real-time interactions [84]. Moreover, their social simulation capabilities create safe, controlled environments for autistic children to practice social skills, generating unlimited possible social scenarios through their generative abilities. This aligns with previous research demonstrating that virtual environments provide secure spaces for children to rehearse social behaviors without real-world unpredictability [86]. By generating dynamic, contextually relevant dialogues, LLMs enhance these virtual environments, enabling practice across diverse social scenarios. While direct real-world experience remains essential, LLMs serve as valuable preparatory tools. The ability to simulate social interactions in virtual settings facilitates repeated practice—crucial for autistic children who often struggle with skill generalization. Research has shown that children who engage with virtual environments for social skills training successfully transfer these skills to real-world contexts [86, 99]. Regarding limitations, hallucination remains one of the most significant risks for LLMs. In our evaluation of 100 simulated responses, we identified 6 hallucination cases (6%): 2 instances of extraneous text generation and 4 cases of intention-behavior inconsistencies. While Unity's processing system filters extraneous text, intention-behavior contradictions are addressed by introducing the concept of "lying cars" to autistic children during experiments. Although the 6% hallucination rate is relatively low, future research could employ self-reflection techniques to further reduce these occurrences [64]. Given the rapid advancement of large language models, subsequent work could explore more advanced models to enable more immediate and diverse interactions.

### 6.2 Benefits and Experiences of Immersive Interfaces for Autistic Children

Projection-based immersive spaces demonstrate significant advantages in social simulation for autistic children, with our medium system proving both accessible and acceptable to them. Embodied learning in spatial environments provides children with enhanced immersion and improved emotional engagement, addressing attention deficits and learning challenges common among autistic children [87]. Their progressively improving performance demonstrates the effectiveness of learning social affordance in virtual scenarios. Simulation serves as a vital bridge before real traffic exposure for autistic children who typically find direct real-world learning challenging - a proven method in special education [110]. While training occurs in virtual spaces, our ultimate goal is to enable autistic children to perform better in real-world situations through these virtual interactions. AIRoad shows promising real-world skill transfer potential, as evidenced by improved response

times to real road footage. The decision-making processes of autistic children in the immersive AIRoad environment closely mirrored those observed in video-based tests. For example, Participant 3 displayed consistent risky behavior, attempting to run across the road during both AIRoad training and the video-based test. Conversely, Participant 1 demonstrated cautious behavior, consistently waiting for vehicles to pass completely before crossing. These observations suggest that immersive environments can effectively expose the real challenges faced by autistic children in a safe context. Given that more accessible interactive projection systems have been developed, further exploration of rehabilitation effects under simpler equipment conditions is possible [36].

### 6.3 Attention Pattern Changes in Autistic Children After AIRoad Training

One possible explanation for the changes in response times lies in the shifts of attention patterns, as revealed by our eye-tracking data during video response tests. After the training, autistic children not only paid attention to traffic lights but also began to observe vehicles and how they followed traffic rules. Prior to AIRoad training, most children understood only the explicit rules for crossing the street—such as stopping at a red light, walking on a green light, and waiting for cars to pass before proceeding. However, these rules often fall short in traffic situations that are rich with complex social signals. The training resulted in a noticeable shift in children's attention patterns towards vehicle signals. This shift is evidenced in eye movement patterns, as shown in Fig. 11. The data reveals that post-training, children allocated more time observing vehicles. While participants initially focused primarily on traffic lights and road conditions, after exploring AIRoad, they redistributed their attention to include parked vehicles. However, this modified attention pattern occasionally led to misjudgments. For example, P11, who previously relied on traffic lights for crossing decisions, began focusing more on vehicle movements post-training, assuming it was safe to cross whenever cars stopped. As P11 noted, *"In the future, I will pay more attention when crossing the road, and I will look out for the driver's hand gestures in the cars."* It is necessary to generate more diverse scenarios to help autistic children integrate multiple cues in their decision-making process. Nevertheless, based on the video-based experiment results, this enriched attention pattern provided children with more comprehensive information for assessing traffic situations.

## 7 Limitations and Future Work

Currently, AIRoad still faces numerous limitations. For example, the hallucinations of LLMs mentioned above are difficult to control [149], leading to the generation of unexpected content in speech and the simulation of affordances, thereby increasing the system's unpredictability. Additionally, the latency of LLMs to some extent affects the system's response time. Our system currently mitigates the impact of delays by generating multiple vehicles simultaneously, but future upgrades and miniaturization of LLMs may potentially improve this issue [124].

Future work could expand to constructing more social scenarios filled with complex signals to simulate their social affordances, including enhanced traffic scenarios with public transportation



**Figure 11: Case on the Changes in Eye Movement Before and After Training with AIRoad. a) Eye movement case before AIRoad training b) Eye movement case after AIRoad training.**

and indoor environments rich with social signals. As Loveland (1991) noted, social affordances are inevitable wherever human environments exist [82]. While the current system primarily relies on positional information as input, future iterations could incorporate additional interaction modalities such as vocal input and tactile interaction. Research suggests that a more diverse range of interaction modalities could enhance children's sense of immersion and lead to more effective learning outcomes [4]. Additionally, exploring the deployment of such educational technologies across broader contexts and on simpler devices represents another promising direction for future research.

## 8 Conclusion

This study began with a comprehensive review of educational technologies for autistic children, identifying 17 design guidelines across four major categories that formed the foundation of our system. By leveraging LLM's social simulation capabilities, we developed an immersive projection-based educational technology that generates rich social affordances. The system uses a street-crossing scenario to simulate diverse driving behaviors, conveying social affordances through multimodal signals including visual cues and auditory prompts. This approach provides autistic children with a safe environment to practice interpreting and responding to social signals. A user study with 14 autistic children revealed significant improvements in their decision-making processes, demonstrating enhanced understanding of social affordances in traffic settings while confirming positive usability and emotional experience. Through a controlled yet dynamic environment, our system effectively supports autistic children in recognizing and responding to complex social cues in traffic settings.

## Acknowledgments

We would like to thank Shanghai Children's Welfare Institute for their inspiration and assistance in this research. This work was supported by the National Natural Science Foundation of China Youth Fund (No. 62202267).

## References

- [1] Frances E Aboud and Maria Amato. 2003. Developmental and socialization influences on intergroup bias. *Blackwell handbook of social psychology: Intergroup processes* (2003), 65–85.
- [2] Ali Adjourlu and Stefania Serafin. 2020. Co-designing a head-mounted display based virtual reality game to teach street-crossing skills to children diagnosed with autism spectrum disorder. In *Interactivity, Game Creation, Design, Learning, and Innovation: 8th EAI International Conference, ArtsIT 2019, and 4th EAI International Conference, DLI 2019, Aalborg, Denmark, November 6–8, 2019, Proceedings*. Springer, 397–405.
- [3] Camila Alvarado, Roberto Munoz, Rodolfo Villarroel, Oscar Acuña, Thiago S Barcelos, and Carlos Becerra. 2017. ValpoDijo: Developing a software that supports the teaching of chilean idioms to children with autism spectrum disorders. In *2017 Twelfth Latin American Conference on Learning Technologies (LACLO)*. IEEE, 1–4.
- [4] Stamatina Anastopoulou. 2004. *Investigating multimodal interactions for the design of learning environments: a case study in science learning*. Ph.D. Dissertation. University of Birmingham.
- [5] Anthony Attwood, Utte Frith, and Beate Hermelin. 1988. The understanding and use of interpersonal gestures by autistic and Down's syndrome children. *Journal of autism and developmental disorders* 18, 2 (1988), 241–257.
- [6] Tony Attwood. 2000. Strategies for improving the social integration of children with Asperger syndrome. *Autism* 4, 1 (2000), 85–100.
- [7] Aaron Bangor, Philip T Kortum, and James T Miller. 2008. An empirical evaluation of the system usability scale. *Intl. Journal of Human–Computer Interaction* 24, 6 (2008), 574–594.
- [8] Mary Barry and Ian Pitt. 2006. Interaction design: a multidimensional approach for learners with autism. In *Proceedings of the 2006 conference on Interaction design and children*. 33–36.
- [9] Laura Bartoli, Franca Garzotto, Mirko Gelsomini, Luigi Oliveto, and Matteo Valoriani. 2014. Designing and evaluating touchless playful interaction for ASD children. In *Proceedings of the 2014 conference on Interaction design and children*. 17–26.
- [10] Renae Beaumont and Kate Sofronoff. 2013. multimodal intervention for social skills Training in students with high-functioning asd. *CBT for children and adolescents with high-functioning autism spectrum disorders* (2013), 173.
- [11] Rongqi Bei, Yajie Liu, Yihe Wang, Yuxuan Huang, Ming Li, Yuhang Zhao, and Xin Tong. 2024. StarRescue: the Design and Evaluation of A Turn-Taking Collaborative Game for Facilitating Autistic Children's Social Skills. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 1–19.
- [12] Esubalew Bekele, Joshua Wade, Dayi Bian, Jing Fan, Amy Swanson, Zachary Warren, and Nilanjan Sarkar. 2016. Multimodal adaptive social interaction in virtual environment (MASI-VR) for children with Autism spectrum disorders (ASD). In *2016 IEEE virtual reality (VR)*. IEEE, 121–130.
- [13] Scott Bellini, Jessica K Peters, Lauren Benner, and Andrea Hopf. 2007. A meta-analysis of school-based social skills interventions for children with autism spectrum disorders. *Remedial and Special Education* 28, 3 (2007), 153–162.
- [14] Sara Bernardini, Kaška Porayska-Pomsta, and Tim J Smith. 2014. ECHOES: An intelligent serious game for fostering social communication in children with autism. *Information Sciences* 264 (2014), 41–60.
- [15] Lal Bozgeyikli, Andrew Raij, Srinivas Katkoori, and Redwan Alqasemi. 2017. A survey on virtual reality for individuals with autism spectrum disorder: design considerations. *IEEE Transactions on Learning Technologies* 11, 2 (2017), 133–151.

- [16] Nancy C Brady, Holly L Storkel, Paige Bushnell, R Michael Barker, Kate Saunders, Debby Daniels, and Kandace Fleming. 2015. Investigating a multimodal intervention for children with limited expressive vocabularies associated with autism. *American Journal of Speech-Language Pathology* 24, 3 (2015), 438–459.
- [17] John Brooke et al. 1996. SUS-A quick and dirty usability scale. *Usability evaluation in industry* 189, 194 (1996), 4–7.
- [18] Patricia J Brooks and Bertram O Ploog. 2013. Attention to emotional tone of voice in speech perception in children with autism. *Research in Autism Spectrum Disorders* 7, 7 (2013), 845–857.
- [19] Tom B. Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, Sandhini Agarwal, Ariel Herbert-Voss, Gretchen Krueger, Tom Henighan, Rewon Child, Aditya Ramesh, Daniel M. Ziegler, Jeffrey Wu, Clemens Winter, Christopher Hesse, Mark Chen, Eric Sigler, Mateusz Litwin, Scott Gray, Benjamin Chess, Jack Clark, Christopher Berner, Sam McCandlish, Alec Radford, Ilya Sutskever, and Dario Amodei. 2020. Language Models are Few-Shot Learners. arXiv:2005.14165 [cs.CL]
- [20] Grigore C Burdea and Philippe Coiffet. 2003. *Virtual reality technology*. John Wiley & Sons.
- [21] Yu Cai, Zhao Liu, Zhuo Yang, Yilan Tan, Junwei Zhang, and Shuo Tang. 2023. Starrypia: An AR Gamified Music Adjuvant Treatment Application for Children with Autism Based on Combined Therapy. In *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology*. 1–16.
- [22] Brianna E Cairney, Stanley H West, Eileen Haebig, Christopher R Cox, and Heather D Lucas. 2023. Interpretations of meaningful and ambiguous hand gestures in autistic and non-autistic adults: A norming study. *Behavior Research Methods* (2023), 1–14.
- [23] Murilo C Camargo, Rodolfo M Barros, Jacques D Brancher, Vanessa TO Barros, and Matheus Santana. 2019. Designing gamified interventions for autism spectrum disorder: a systematic review. In *Entertainment Computing and Serious Games: First IFIP TC 14 Joint International Conference, ICEC-JCSG 2019, Arequipa, Peru, November 11–15, 2019, Proceedings* 1. Springer, 341–352.
- [24] Alice S Carter, Naomi Ornstein Davis, Ami Klin, and Fred R Volkmar. 2005. Social development in autism. *Handbook of autism and pervasive developmental disorders* 1 (2005), 312–334.
- [25] Erik W Carter, Eric A Common, Melissa A Sreckovic, Heartley B Huber, Kristen Bottema-Beutel, Jenny Redding Gustafson, Jessica Dykstra, and Kara Hume. 2014. Promoting social competence and peer relationships for adolescents with autism spectrum disorders. *Remedial and Special Education* 35, 2 (2014), 91–101.
- [26] Giorgio Celani, Marco Walter Battacchi, and Letizia Arcidiacono. 1999. The understanding of the emotional meaning of facial expressions in people with autism. *Journal of autism and developmental disorders* 29 (1999), 57–66.
- [27] Ismail Celik, Muhterem Dindar, Hanni Muukkonen, and Sanna Järvelä. 2022. The Promises and Challenges of Artificial Intelligence for Teachers: a Systematic Review of Research. *TechTrends* 66, 4 (01 Jul 2022), 616–630. <https://doi.org/10.1007/s11528-022-00715-y>
- [28] Inha Cha, Sung-In Kim, Hwajung Hong, Heejeong Yoo, and Youn-kyung Lim. 2021. Exploring the use of a voice-based conversational agent to empower adolescents with autism spectrum disorder. In *Proceedings of the 2021 CHI conference on human factors in computing systems*. 1–15.
- [29] Brandt Chamberlain, Connie Kasari, and Erin Rotheram-Fuller. 2007. Involvement or isolation? The social networks of children with autism in regular classrooms. *Journal of autism and developmental disorders* 37 (2007), 230–242.
- [30] Konstantina Chatzara, Charalampos Karagiannidis, Sofia Mavropoulou, and Dimos Stamatis. 2014. Digital storytelling for children with autism: software development and pilot application. In *Research on e-learning and ICT in education: Technological, pedagogical and instructional perspectives*. Springer, 287–300.
- [31] Ya-Yun Chen, Mirko Uljarevic, Joshua Neal, Steven Greening, Hyungwook Yim, and Tae-Ho Lee. 2022. Excessive functional coupling with less variability between salience and default mode networks in autism spectrum disorder. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging* 7, 9 (2022), 876–884.
- [32] Yufang Cheng, Cheng-Li Huang, and Chung-Sung Yang. 2015. Using a 3D immersive virtual environment system to enhance social understanding and social skills for children with autism spectrum disorders. *Focus on Autism and Other Developmental Disabilities* 30, 4 (2015), 222–236.
- [33] Yujin Cho, Mingeon Kim, Seojin Kim, Oyun Kwon, Ryan Donghan Kwon, Yoonha Lee, and Dohyun Lim. 2023. Evaluating the efficacy of interactive language therapy based on llm for high-functioning autistic adolescent psychological counseling. *arXiv preprint arXiv:2311.09243* (2023).
- [34] Lennard Chua, Jeremy Goh, Zin Tun Nay, Lihui Huang, Yiyu Cai, and Ruby Seah. 2017. ICT-enabled emotional learning for special needs education. *Simulation and Serious Games for Education* (2017), 29–45.
- [35] Mauro Conti, Roberto Di Pietro, Luigi V. Mancini, and Alessandro Mei. 2009. (old) Distributed data source verification in wireless sensor networks. *Inf. Fusion* 10, 4 (2009), 342–353. <https://doi.org/10.1016/j.inffus.2009.01.002>
- [36] Giulia Cosentino, Mirko Gelsomini, and Michail Giannakos. 2023. MOVES: Going beyond hardwired multisensory environments for children. In *Proceedings of the 22nd Annual ACM Interaction Design and Children Conference*. 716–720.
- [37] Allison E Curry, Kristina B Metzger, Meghan E Carey, Emma B Sartin, Patty Huang, and Benjamin E Yerys. 2021. Comparison of motor vehicle crashes, traffic violations, and license suspensions between autistic and non-autistic adolescent and young adult drivers. *Journal of the American Academy of Child & Adolescent Psychiatry* 60, 7 (2021), 913–923.
- [38] Kerstin Dautenhahn. 2000. Design issues on interactive environments for children with autism. In *Proc of ICDVRAT 2000, the 3rd Int Conf on Disability, Virtual Reality and Associated Technologies*. University of Reading.
- [39] Megan Davis, Kerstin Dautenhahn, Stuart Powell, and Chrysopher Nehaniv. 2010. Guidelines for researchers and practitioners designing software and software trials for children with autism. *Journal of Assistive Technologies* 4, 1 (2010), 38–48.
- [40] Vanessa Tavares de Oliveira Barros, Cristiane Affonso de Almeida Zerbetto, Kátia Tavares Meserian, Rodolfo Barros, Murilo Crivellari Camargo, and Táthia Cristina Passos de Carvalho. 2014. DayByDay: Interactive and customizable use of mobile technology in the cognitive development process of children with autistic spectrum disorder. In *Universal Access in Human-Computer Interaction. Universal Access to Information and Knowledge: 8th International Conference, UAHCI 2014, Held as Part of HCI International 2014, Heraklion, Crete, Greece, June 22–27, 2014, Proceedings, Part II* 8. Springer, 443–453.
- [41] Marc Fabri and David Moore. 2005. The use of emotionally expressive avatars in collaborative virtual environments. *Virtual Social Agents* 88, April (2005), 88–94.
- [42] Torbjörn Falkmer, Anna Anund, Gunilla Sörensen, and Marita Falkmer. 2004. The transport mobility situation for children with autism spectrum disorders. *Scandinavian Journal of Occupational Therapy* 11, 2 (2004), 90–100.
- [43] Megan Freeth, Tom Foulsham, and Alan Kingstone. 2013. What affects social attention? Social presence, eye contact and autistic traits. *PLoS one* 8, 1 (2013), e53286.
- [44] Uta Frith. 1994. Autism and theory of mind in everyday life. *Social development* 3, 2 (1994), 108–124.
- [45] Chen Gao, Xiaochong Lan, Zhihong Lu, Jinzhu Mao, Jinghua Piao, Huadong Wang, Depeng Jin, and Yong Li. 2023. S3: Social-network Simulation System with Large Language Model-Empowered Agents. arXiv:2307.14984 [cs.SI]
- [46] James G Gibson. 1977. The theory of affordances. *Hilldale, USA* 1, 2 (1977), 67–82.
- [47] Tina R Goldsmith. 2008. *Using virtual reality enhanced behavioral skills training to teach street-crossing skills to children and adolescents with Autism Spectrum Disorders*. Western Michigan University.
- [48] Jiangtao Gong, Teng Han, Siling Guo, Jiannan Li, Siyu Zha, Liuxin Zhang, Feng Tian, Qianying Wang, and Yong Rui. 2021. Holoboard: A large-format immersive teaching board based on pseudo holographics. In *The 34th Annual ACM Symposium on User Interface Software and Technology*. 441–456.
- [49] Marine Grandgeorge and Nobuo Masataka. 2016. Atypical color preference in children with autism spectrum disorder. *Frontiers in psychology* 7 (2016), 227653.
- [50] Temple Grandin. 2002. 28 Teaching Tips for Children and Adults with Autism. (2002).
- [51] Ouriel Grynszpan, Patrice L Weiss, Fernando Perez-Diaz, and Eynat Gal. 2014. Innovative technology-based interventions for autism spectrum disorders: a meta-analysis. *Autism* 18, 4 (2014), 346–361.
- [52] Ge Guo and Zhiwen Qiu. 2024. DentAR: Innovating Dental Visits with Sensory Experiences in AR for People with Autism Spectrum Disorder. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*. 1–7.
- [53] Osama Halabi, Samir A Elseoud, Jihad M Alja'am, Hena Alpona, Moza Al-Hemadi, and Dabia Al-Hassan. 2017. Immersive virtual reality in improving communication skills in children with autism. (2017).
- [54] Joanna Hale, Jamie A Ward, Francesco Buccheri, Dominic Oliver, and Antonia F de C Hamilton. 2020. Are you on my wavelength? Interpersonal coordination in dyadic conversations. *Journal of nonverbal behavior* 44 (2020), 63–83.
- [55] Annika Hellendoorn. 2014. Understanding social engagement in autism: being different in perceiving and sharing affordances. *Frontiers in Psychology* 5 (2014), 98969.
- [56] Molly Helt, Elizabeth Kelley, Marcel Kinsbourne, Juhi Pandey, Hilary Boorstein, Martha Herbert, and Deborah Fein. 2008. Can children with autism recover? If so, how? *Neuropsychology review* 18 (2008), 339–366.
- [57] R Peter Hobson. 1989. On sharing experiences. *Development and psychopathology* 1, 3 (1989), 197–203.
- [58] R Peter Hobson. 1990. On acquiring knowledge about people and the capacity to pretend: response to Leslie (1987). (1990).
- [59] R Peter Hobson. 1991. Against the theory of ‘theory of mind’. *British Journal of Developmental Psychology* 9, 1 (1991), 33–51.
- [60] R Peter Hobson. 1993. The emotional origins of social understanding. *Philosophical psychology* 6, 3 (1993), 227–249.
- [61] Juan Pablo Hourcade, Natasha E Bullock-Rest, and Thomas E Hansen. 2012. Multitouch tablet applications and activities to enhance the social skills of children with autism spectrum disorders. *Personal and ubiquitous computing* 16

- (2012), 157–168.
- [62] Che-Wei Hsu and Yee-San Teoh. 2017. Investigating event memory in children with autism spectrum disorder: Effects of a computer-mediated interview. *Journal of autism and developmental disorders* 47 (2017), 359–372.
- [63] Kara Hume, Rachel Loftin, and Johanna Lantz. 2009. Increasing independence in autism spectrum disorders: A review of three focused interventions. *Journal of autism and developmental disorders* 39 (2009), 1329–1338.
- [64] Ziwei Ji, Tiezheng Yu, Yan Xu, Nayeon Lee, Etsuko Ishii, and Pascale Fung. 2023. Towards mitigating LLM hallucination via self reflection. In *Findings of the Association for Computational Linguistics: EMNLP 2023*. 1827–1843.
- [65] Naomi Josman, Hadass Milik Ben-Chaim, Shula Friedrich, and Patrice L Weiss. 2008. Effectiveness of virtual reality for teaching street-crossing skills to children and adolescents with autism. *International Journal on Disability and Human Development* 7, 1 (2008), 49–56.
- [66] Leo Kanner, Alejandro Rodriguez, and Barbara Ashenden. 1972. How far can autistic children go in matters of social adaptation? *Journal of autism and childhood schizophrenia* 2 (1972), 9–33.
- [67] Connie Kasari, Jill Locke, Amanda Gulsrud, and Erin Rotheram-Fuller. 2011. Social networks and friendships at school: Comparing children with and without ASD. *Journal of autism and developmental disorders* 41 (2011), 533–544.
- [68] Fengfeng Ke and Sungwoong Lee. 2016. Virtual reality based collaborative design by children with high-functioning autism: Design-based flexibility, identity, and norm construction. *Interactive Learning Environments* 24, 7 (2016), 1511–1533.
- [69] Norman Kiak Nam Kee, Noel Kok Hwee Chia, and Yiyu Cai. 2012. Universal design for learning (UDL) and living (UDL2) in virtual reality-based treatments for children with autism. In *Proceedings of the 6th International Conference on Rehabilitation Engineering & Assistive Technology*. 1–4.
- [70] Robert L Koegel, Lynn Kern Koegel, and Lauren I Brookman. 2003. Empirically supported pivotal response interventions for children with autism. *Evidence-based psychotherapies for children and adolescents* 34(357) (2003).
- [71] Evdokimos I Konstantinidis, Magda Hitoglou-Antoniadou, Andrej Luneski, Panagiotis D Bamidis, and Maria M Nikolaidou. 2009. Using affective avatars and rich multimedia content for education of children with autism. In *Proceedings of the 2nd international conference on pervasive technologies related to assistive environments*. 1–6.
- [72] Rahadian Kurniawan, Wuriandiety Mayang Purnamasari, Restu Rakhmawati, and Dimas Panji Eka Jalaputra. 2018. Development of game for self-help toilet learning for children with Autism. *CommIT (Communication and Information Technology) Journal* 12, 1 (2018), 1–12.
- [73] Cecilia Sik Lányi and Ádám Tilinger. 2004. Multimedia and virtual reality in the rehabilitation of autistic children. In *International Conference on Computers for Handicapped Persons*. Springer, 22–28.
- [74] Wan-Ju Lee, Chi-Wen Huang, Chia-Jung Wu, Shing-Tsaan Huang, and Gwo-Dong Chen. 2012. The effects of using embodied interactions to improve learning performance. In *2012 IEEE 12th international conference on advanced learning technologies*. IEEE, 557–559.
- [75] Yuan Li, Yixuan Zhang, and Lichao Sun. 2023. MetaAgents: Simulating Interactions of Human Behaviors for LLM-based Task-oriented Coordination via Collaborative Generative Agents. *arXiv:2310.06500 [cs.AI]*
- [76] Lili Liu, Xinyu Yao, Jingying Chen, Kun Zhang, Leyuan Liu, Guangshuai Wang, and Yutao Ling. 2024. Virtual reality utilized for safety skills training for autistic individuals: a review. *Behavioral Sciences* 14, 2 (2024), 82.
- [77] Ruibo Liu, Ruixin Yang, Chenyan Jia, Ge Zhang, Denny Zhou, Andrew M. Dai, Dify Yang, and Soroush Vosoughi. 2023. Training Socially Aligned Language Models on Simulated Social Interactions. *arXiv:2305.16960 [cs.CL]*
- [78] Xiongyi Liu, Qing Wu, Wenbing Zhao, and Xiong Luo. 2017. Technology-facilitated diagnosis and treatment of individuals with autism spectrum disorder: An engineering perspective. *Applied Sciences* 7, 10 (2017), 1051.
- [79] Pedro Lopes, Patrik Jonell, and Patrick Baudisch. 2015. Affordance++ allowing objects to communicate dynamic use. In *Proceedings of the 33rd annual ACM conference on human factors in computing systems*. 2515–2524.
- [80] Timo Lorenz, Nomi Reznik, and Kathrin Heinitz. 2017. A different point of view: The neurodiversity approach to autism and work. *Autism: Paradigms, Recent Research, and Clinical Applications* (2017).
- [81] Gonzalo Lorenzo, Asunción Lledó, Jorge Pomares, and Rosabel Roig. 2016. Design and application of an immersive virtual reality system to enhance emotional skills for children with autism spectrum disorders. *Computers & Education* 98 (2016), 192–205.
- [82] Katherine A Loveland. 1991. Social affordances and interaction II: Autism and the affordances of the human environment. *Ecological psychology* 3, 2 (1991), 99–119.
- [83] Shame L Lynch and Angela N Irvine. 2009. Inclusive education and best practice for children with autism spectrum disorder: An integrated approach. *International journal of inclusive education* 13, 8 (2009), 845–859.
- [84] Yue Lyu, Pengcheng An, Huan Zhang, Keiko Katsuragawa, and Jian Zhao. 2024. Designing AI-Enabled Games to Support Social-Emotional Learning for Children with Autism Spectrum Disorders. *arXiv preprint arXiv:2404.15576* (2024).
- [85] Matthew J Maenner. 2023. Prevalence and characteristics of autism spectrum disorder among children aged 8 years—Autism and Developmental Disabilities Monitoring Network, 11 sites, United States, 2020. *MMWR. Surveillance Summaries* 72 (2023).
- [86] Skevi Matsentidou and Charalambos Poullis. 2014. Immersive visualizations in a VR cave environment for the training and enhancement of social skills for children with autism. In *2014 international conference on computer vision theory and applications (VISAPP)*, Vol. 3. IEEE, 230–236.
- [87] James P McGee and Catherine Lord. 2001. *Educating children with autism*. National Academies Press.
- [88] Chan Mei, Lee Mason, and John Quarles. 2014. Usability issues with 3D user interfaces for adolescents with high functioning autism. In *Proceedings of the 16th international ACM SIGACCESS conference on Computers & accessibility*. 99–106.
- [89] Laura Millen, Rob Edlin-White, and Sue Cobb. 2010. The development of educational collaborative virtual environments for children with autism. In *Proceedings of the 5th Cambridge Workshop on Universal Access and Assistive Technology, Cambridge*, Vol. 1.
- [90] Ruchir Mishra and Karla Conn Welch. 2024. Towards scalable robotic intervention of children with Autism Spectrum Disorder using LLMs. *arXiv preprint arXiv:2402.00260* (2024).
- [91] Fereshteh Mohammadzaheri, Lynn Kern Koegel, Mohammad Rezaee, and Seyed Majid Rafiee. 2014. A randomized clinical trial comparison between pivotal response treatment (PRT) and structured applied behavior analysis (ABA) intervention for children with autism. *Journal of autism and developmental disorders* 44 (2014), 2769–2777.
- [92] Maggie A Mosher, Adam C Carreon, Stephanie L Craig, and Lindsay C Ruhter. 2022. Immersive technology to teach social skills to students with autism spectrum disorder: A literature review. *Review Journal of Autism and Developmental Disorders* 9, 3 (2022), 334–350.
- [93] Rebecca Muhle, Stephanie V Trentacoste, and Isabelle Rapin. 2004. The genetics of autism. *Pediatrics* 113, 5 (2004), e472–e486.
- [94] H Neale, A Leonard, and S Kerr. 2002. Exploring the role of virtual environments in the special needs classroom. In *Proceedings of the 4th ICDVRAT*. Veszprem Hungary, 259–266.
- [95] Don Norman. 2013. *The design of everyday things: Revised and expanded edition*. Basic books.
- [96] Donald A Norman. 1988. *The psychology of everyday things*. Basic books.
- [97] Mmachi God'sglory Obiorah, Anne Marie Marie Piper, and Michael Horn. 2021. Designing AACs for people with aphasia dining in restaurants. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–14.
- [98] François Osiurak, Yves Rossetti, and Arnaud Badets. 2017. What is an affordance? 40 years later. *Neuroscience & Biobehavioral Reviews* 77 (2017), 403–417.
- [99] Biplov Paneru and Bushwash Paneru. 2024. The Nexus of AR/VR, Large Language Models, UI/UX, and Robotics Technologies in Enhancing Learning and Social Interaction for Children: A Systematic Review. *arXiv preprint arXiv:2409.18162* (2024).
- [100] Xianghe Pang, Shuo Tang, Rui Ye, Yuxin Xiong, Bolun Zhang, Yanfeng Wang, and Siheng Chen. 2024. Self-alignment of large language models via multi-agent social simulation. In *ICLR 2024 Workshop on Large Language Model (LLM) Agents*.
- [101] Joon Sung Park, Joseph O'Brien, Carrie Jun Cai, Meredith Ringel Morris, Percy Liang, and Michael S Bernstein. 2023. Generative agents: Interactive simulators of human behavior. In *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology*. 1–22.
- [102] Sarah Parsons and Sue Cobb. 2013. Who chooses what I need? Child voice and user-involvement in the development of learning technologies for children with autism. *EPSRC Observatory for Responsible Innovation in ICT* (2013).
- [103] Sarah Parsons, Anne Leonard, and Peter Mitchell. 2006. Virtual environments for social skills training: comments from two adolescents with autistic spectrum disorder. *Computers & Education* 47, 2 (2006), 186–206.
- [104] Sarah Parsons, Peter Mitchell, and Anne Leonard. 2005. Do adolescents with autistic spectrum disorders adhere to social conventions in virtual environments? *Autism* 9, 1 (2005), 95–117.
- [105] Marie Pelé, Caroline Bellut, Elise Debergue, Charlotte Gauvin, Anne Jeanneret, Thibault Leclerc, Lucie Nicolas, Florence Pontier, Diorne Zausa, and Cédric Sueur. 2017. Cultural influence of social information use in pedestrian road-crossing behaviours. *Royal Society open science* 4, 2 (2017), 160739.
- [106] Yicong Peng, Wei Zhu, Fangyu Shi, Yi Fang, and Guangtao Zhai. 2019. Virtual reality based road crossing training for autistic children with behavioral analysis. In *Digital TV and Multimedia Communication: 15th International Forum, IFTC 2018, Shanghai, China, September 20–21, 2018, Revised Selected Papers 15*. Springer, 456–469.
- [107] Kaska Porayska-Pomsta, Christopher Frauenberger, Helen Pain, Gnanathusharan Rajendran, Tim Smith, Rachel Menzies, Mary Ellen Foster, Alyssa Alcorn, Sam Wass, Sara Bernadini, et al. 2012. Developing technology for autism: an interdisciplinary approach. *Personal and Ubiquitous Computing* 16 (2012), 117–127.

- [108] Kathleen A Quill. 1997. Instructional considerations for young children with autism: The rationale for visually cued instruction. *Journal of autism and developmental disorders* 27, 6 (1997), 697–714.
- [109] Maxwell JD Ramstead, Samuel PL Veissière, and Laurence J Kirmayer. 2016. Cultural affordances: Scaffolding local worlds through shared intentionality and regimes of attention. *Frontiers in psychology* 7 (2016), 206814.
- [110] Patricia A Rao, Deborah C Beidel, and Michael J Murray. 2008. Social skills interventions for children with Asperger's syndrome or high-functioning autism: A review and recommendations. *Journal of autism and developmental disorders* 38 (2008), 353–361.
- [111] François Robain, Michel Godel, Nada Kojovic, Martina Franchini, Fiona Journal, and Marie Schaer. 2022. Measuring social orienting in preschoolers with autism spectrum disorder using cartoons stimuli. *Journal of Psychiatric Research* 156 (2022), 398–405.
- [112] Delphine B Rosset, Cecile Rondan, David Da Fonseca, Andreia Santos, Brigitte Assouline, and Christine Deruelle. 2008. Typical emotion processing for cartoon but not for real faces in children with autistic spectrum disorders. *Journal of autism and developmental disorders* 38 (2008), 919–925.
- [113] Erin Rotheram-Fuller, Connie Kasari, Brandt Chamberlain, and Jill Locke. 2010. Social involvement of children with autism spectrum disorders in elementary school classrooms. *Journal of Child Psychology and Psychiatry* 51, 11 (2010), 1227–1234.
- [114] James A Russell. 1980. A circumplex model of affect. *Journal of personality and social psychology* 39, 6 (1980), 1161.
- [115] James A Russell, Anna Weiss, and Gerald A Mendelsohn. 1989. Affect grid: a single-item scale of pleasure and arousal. *Journal of personality and social psychology* 57, 3 (1989), 493.
- [116] Mario Saiano, Laura Pellegrino, Maura Casadio, Susanna Summa, Eleonora Garbarino, Valentina Rossi, Daniela Dall'Agata, and Vittorio Sanguineti. 2015. Natural interfaces and virtual environments for the acquisition of street crossing and path following skills in adults with Autism Spectrum Disorders: a feasibility study. *Journal of neuroengineering and rehabilitation* 12 (2015), 1–13.
- [117] Katayoun Salamatı, Bastian J Schroeder, Duane R Geruschat, and Nagui M Roushail. 2013. Event-based modeling of driver yielding behavior to pedestrians at two-lane roundabout approaches. *Transportation research record* 2389, 1 (2013), 1–11.
- [118] Stefanie Schelinski and Katharina von Kriegstein. 2019. The relation between vocal pitch and vocal emotion recognition abilities in people with autism spectrum disorder and typical development. *Journal of autism and developmental disorders* 49 (2019), 68–82.
- [119] Richard C Schmidt and Michael J Richardson. 2008. Dynamics of interpersonal coordination. In *Coordination: Neural, behavioral and social dynamics*. Springer, 281–308.
- [120] Laura Schreibman and O Ivar Lovaas. 1973. Overselective response to social stimuli by autistic children. *Journal of Abnormal Child Psychology* 1, 2 (1973), 152–168.
- [121] Leila Shahmoradi and Sorayya Rezai. 2022. Cognitive rehabilitation in people with autism spectrum disorder: a systematic review of emerging virtual reality-based approaches. *Journal of NeuroEngineering and Rehabilitation* 19, 1 (2022), 91.
- [122] Murray Shanahan, Kyle McDonell, and Laria Reynolds. 2023. Role play with large language models. *Nature* 623, 7987 (01 Nov 2023), 493–498. <https://doi.org/10.1038/s41586-023-06647-8>
- [123] Yongliang Shen, Kaitao Song, Xu Tan, Dongsheng Li, Weiming Lu, and Yueling Zhuang. 2023. HuggingGPT: Solving AI Tasks with ChatGPT and its Friends in Hugging Face. arXiv:2303.17580 [cs.CL]
- [124] Xuemei Shi, Xiaoguang Du, and Xiedong Song. 2023. Research on Miniaturization Trend of ChatGPT Technology Model. *Journal of Artificial Intelligence and Technology* 3, 3 (2023), 95–99.
- [125] Priyamvada Singh, Rohit Rathore, Rashmi Chauhan, and RH Goudar. 2012. A gaming Model for teaching autistic children basic traffic rules. *International Journal of Information and Education Technology* 2, 5 (2012), 507.
- [126] I Siti Iradah and AK Rabiah. 2011. EduTism: an assistive educational system for the treatment of autism children with intelligent approach. In *Visual Informatics: Sustaining Research and Innovations: Second International Visual Informatics Conference IVIC 2011, Selangor, Malaysia, November 9–11, 2011, Proceedings, Part II 2*. Springer, 193–204.
- [127] Leann E Smith, Matthew J Maenner, and Marsha Mailick Seltzer. 2012. Developmental trajectories in adolescents and adults with autism: The case of daily living skills. *Journal of the American Academy of Child & Adolescent Psychiatry* 51, 6 (2012), 622–631.
- [128] Aiganyra Soltiyeva, Wilk Oliveira, Alimanova Madina, Shygys Adilkhan, Marat Urmanov, and Juho Hamari. 2023. My Lovely Granny's Farm: An immersive virtual reality training system for children with autism spectrum disorder. *Education and Information Technologies* 28, 12 (2023), 16887–16907.
- [129] Dorothy C Strickland, David McAllister, Claire D Coles, and Susan Osborne. 2007. An evolution of virtual reality training designs for children with autism and fetal alcohol spectrum disorders. *Topics in language disorders* 27, 3 (2007), 226–241.
- [130] Teo Susnjak. 2022. ChatGPT: The End of Online Exam Integrity? arXiv:2212.09292 [cs.AI]
- [131] Yilin Tang, Liuqing Chen, Ziyu Chen, Wenkai Chen, Yu Cai, Yao Du, Fan Yang, and Lingyun Sun. 2024. EmoEden: Applying Generative Artificial Intelligence to Emotional Learning for Children with High-Function Autism. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 1–20.
- [132] Orit Taubman-Ben-Ari, Mario Mikulincer, and Omri Gillath. 2004. The multidimensional driving style inventory—scale construct and validation. *Accident Analysis & Prevention* 36, 3 (2004), 323–332.
- [133] Max Thorsson, Martyna A Galazka, Jakob Åberg Johnels, and Nouchine Hadjikhani. 2024. Influence of autistic traits and communication role on eye contact behavior during face-to-face interaction. *Scientific Reports* 14, 1 (2024), 8162.
- [134] Michael Tomasello. 2014. *A natural history of human thinking*. Harvard University Press.
- [135] Wei-Tc Tsai, I-Jui Lee, and Chien-Hsu Chen. 2021. Inclusion of third-person perspective in CAVE-like immersive 3D virtual reality role-playing games for social reciprocity training of children with an autism spectrum disorder. *Universal Access in the Information Society* 20, 2 (2021), 375–389.
- [136] Aimilia Tzanavari, Nefi Charalambous-Darden, Kyriakos Herakleous, and Charalambos Poullis. 2015. Effectiveness of an Immersive Virtual Environment (CAVE) for teaching pedestrian crossing to children with PDD-NOS. In *2015 IEEE 15th international conference on advanced learning technologies*. IEEE, 423–427.
- [137] Nkiruka Uzuegbunam, Wing Hang Wong, Sen-Ching Samson Cheung, and Lisa Ruble. 2017. MEBook: multimedia social greetings intervention for children with autism spectrum disorders. *IEEE Transactions on Learning Technologies* 11, 4 (2017), 520–535.
- [138] Helma Van Rijn, Pieter Jan Stappers, et al. 2008. The puzzling life of autistic toddlers: design guidelines from the LINKX project. *Advances in Human-Computer Interaction* 2008 (2008).
- [139] Koen Vellenga, H Joe Steinhauer, Alexander Karlsson, Göran Falkman, Asli Rhodin, and Ashok Chaitanya Koppiisetty. 2022. Driver intention recognition: State-of-the-art review. *IEEE Open Journal of Intelligent Transportation Systems* 3 (2022), 602–616.
- [140] Diogo Vieira, Ana Leal, Nuno Almeida, Samuel Silva, and António Teixeira. 2017. “Tell Your Day”: Developing Multimodal Interaction Applications for Children with ASD. In *Universal Access in Human–Computer Interaction. Design and Development Approaches and Methods: 11th International Conference, UAHCI 2017, Held as Part of HCI International 2017, Vancouver, BC, Canada, July 9–14, 2017, Proceedings, Part I 11*. Springer, 525–544.
- [141] Susanna Värtinen, Perttu Hämäläinen, and Christian Guckelsberger. 2024. Generating Role-Playing Game Quests With GPT Language Models. *IEEE Transactions on Games* 16, 1 (2024), 127–139. <https://doi.org/10.1109/TG.2022.3228480>
- [142] S Jane Weeks and R Peter Hobson. 1987. The salience of facial expression for autistic children. *Journal of Child Psychology and Psychiatry* 28, 1 (1987), 137–152.
- [143] Elisabeth M Whyte, Joshua M Smyth, and K Suzanne Scherf. 2015. Designing serious game interventions for individuals with autism. *Journal of autism and developmental disorders* 45 (2015), 3820–3831.
- [144] Chathurika S Wickramasinghe, Daniel L Marino, Javier Grandio, and Milos Manic. 2020. Trustworthy AI development guidelines for human system interaction. In *2019 International Conference on Human System Interaction (HSI)*. IEEE, 130–136.
- [145] Connie Wong, Samuel L Odom, Kara A Hume, Ann W Cox, Angel Fettig, Suzanne Kucharczyk, Matthew E Brock, Joshua B Plavnick, Veronica P Fleury, and Tia R Schultz. 2015. Evidence-based practices for children, youth, and young adults with autism spectrum disorder: A comprehensive review. *Journal of autism and developmental disorders* 45 (2015), 1951–1966.
- [146] Rebecca Wood. 2019. *Inclusive education for autistic children: Helping children and young people to learn and flourish in the classroom*. Jessica Kingsley Publishers.
- [147] Qin Wu, Chenmei Yu, Yanjun Chen, Jiayu Yao, Xi Wu, Xiaolan Peng, and Teng Han. 2020. Squeeze the ball: Designing an interactive playground towards aiding social activities of children with low-function autism. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–14.
- [148] Yudan Wu, Shanhe You, Zixuan Guo, Xiangyang Li, Guyue Zhou, and Jiangtao Gong. 2023. MR. Brick: designing a remote mixed-reality educational game system for promoting children's social & collaborative skills. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–18.
- [149] Ziwei Xu, Sanjay Jain, and Mohan Kankanhalli. 2024. Hallucination is Inevitable: An Innate Limitation of Large Language Models. arXiv:2401.11817 [cs.CL]
- [150] Feng Yan. 2011. A sunny day: Ann and ron's world an ipad application for children with autism. In *Serious Games Development and Applications: Second International Conference, SCDA 2011, Lisbon, Portugal, September 19–20, 2011. Proceedings 2*. Springer, 129–138.
- [151] Hui Yang, Sifu Yue, and Yunzhong He. 2023. Auto-GPT for Online Decision Making: Benchmarks and Additional Opinions. arXiv:2306.02224 [cs.AI]

- [152] Aaron Zaffke, Niharika Jain, Norah Johnson, Mohammad Arif Ul Alam, Marta Magiera, and Sheikh Iqbal Ahamed. 2015. iCanLearn: A mobile application for creating flashcards and social stories™ for children with autism. In *Smart Homes and Health Telematics: 12th International Conference, ICOST 2014, Denver, CO, USA, June 25–27, 2014, Revised Papers 12*. Springer, 225–230.
- [153] Irene Zanardi. 2022. Hoomie Sweet Home: A Qualitative Study on the Small Multisensory Space "Hoomie" for Inclusion of Autistic Kids in a Primary School. In *CHI Conference on Human Factors in Computing Systems Extended Abstracts*. 1–4.
- [154] Lian Zhang, Qiang Fu, Amy Swanson, Amy Weitlauf, Zachary Warren, and Nilanjana Sarkar. 2018. Design and evaluation of a collaborative virtual environment (CoMove) for autism spectrum disorder intervention. *ACM Transactions on Accessible Computing (TACCESS)* 11, 2 (2018), 1–22.
- [155] Wenbing Zhao, Xiongyi Liu, Tie Qiu, and Xiong Luo. 2020. Virtual avatar-based life coaching for children with autism spectrum disorder. *Computer* 53, 2 (2020), 26–34.
- [156] Chen Zheng and Barbara Tversky. 2024. Putting it Together, Together. *Cognitive Science* 48, 2 (2024), e13405.
- [157] Cheng Zheng, Caowei Zhang, Xuan Li, Xin Liu, Chuqi Tang, Guanyun Wang, Cheng Yao, Fan Zhang, Wenjie Xu, and Fangtian Ying. 2017. Toon-chat: A cartoon-masked chat system for children with autism. In *ACM SIGGRAPH 2017 Posters*. 1–2.

## A Experimental Details

### A.1 Online questionnaire

**Investigation on ASD children's language and motor skills.**  
Hello, Parents! Thank you very much for participating in this questionnaire. We have designed a virtual space environment for autistic children to explore, where they can learn and interact.  
All your answers are for statistical analysis and academic research only. Please fill in the questionnaire carefully according to the actual situation of your child and feel free to answer! If you and your child are interested in participating in our "Little Explorer" virtual space experience and becoming our little explorers, please leave your contact information at the end of the questionnaire.

Your child's name: What is your child's gender?

- male
- female

Please enter your child's date of birth:

How is your child's motor skills?

- Very good
- Good
- Average
- Poor
- Very poor

How is your child's language skills?

- Very good
- Good
- Average
- Poor
- Very poor

Your child's personality:

- Outgoing and lively
- Introverted and quiet
- Curious and exploratory
- Sensitive and sentimental

Does your child enjoy going on trips?

- Like very much
- Not so much
- Neutral

- Dislike
- Very Dislike

What is your usual mode of trips?

- Self-drive
- Public Transport
- Walking
- Cycling
- Other

Has your child travelled independently?

- Yes, he/she can travel fully independently
- limited independent trips, need some guidance
- cannot travel independently, always needs to be accompanied
- never tried to travel independently

## B Adapted SUS Scale

(five-Likert scale, 1:It fits perfectly;5:Very inconsistent)

### B.1 AIRoad

- I think I would let my child play this game often
- I think this game is too difficult for children
- I think this game is too easy for the child to play
- The child needs help to complete the game
- I think the system is well integrated with various functions
- I think other children would learn the game very quickly as well
- I think there are too many inconsistencies in the system
- I found the system very cumbersome to use
- I think the system is well designed and fun for children to play
- The child would need to know a lot of other knowledge and information beforehand to play the game

### B.2 Video-tutorial

- I think I would let my child watch these videos often
- I think this video is too difficult for children
- I think this video is too easy for my child to watch
- My child needs help to watch this video independently
- I think this video has a good integration of various elements
- I think other children would learn the content of this video very quickly as well
- I think there are too many inconsistencies in the video
- I found it very cumbersome to watch and use the video
- I think the video is well designed and interesting for children to watch
- Children need to know a lot of other knowledge and information before they can watch this video