Teaching Firearm Safety Behaviors to Young Children Using Robots and Augmented Reality

Maximilian von Kahle[1], Abby Goss[1], Payton Przyuski[1], Justin Kim[1],
Sujan Sarker[1], Bobby Doyle[2], and Tariq Iqbal[1*]
**Corresponding Author: tiqbal@virginia.edu*
[1]School of Engineering and Applied Science, University of Virginia, Charlottesville, VA
[2]Gun Violence Solutions Project, University of Virginia, Charlottesville, VA

***Abstract*—Firearm-related injuries remain a leading cause of death among children in the United States. Behavioral skills training (BST) is widely regarded as the gold standard for teaching children the "Stop, Don't Touch, Run Away, Tell a Trusted Adult" response; however, its conventional delivery relies on trained facilitators and controlled settings, limiting scalability. Despite advances in artificial intelligence, augmented reality, and social robotics, integrated AI-based systems for behaviorally grounded firearm safety education in young children remain underexplored. To address this gap, we propose an AI-based firearm safety training system with two training modules, a robot-based module and an augmented reality (AR)-based module, and a separate AR-based testing module. The system employs a large language model-driven dialogue framework to deliver evidence-based instruction to children aged 4- 7, grounded in domain knowledge via a retrieval-augmented architecture. It comprises three modules aligned with BST stages: robot-delivered instruction, AR-guided behavioral rehearsal, and AR-based in-situ evaluation for skill generalization. We evaluate the system through expert review with professionals in child development, education, and injury prevention. Expert feedback indicates that the system is a promising approach for safety-critical behavioral training, with strengths in usability, engagement, and emotionally appropriate instruction, while highlighting areas for refinement, including language accessibility for younger children, trauma-informed interaction strategies, consistency in AI-driven dialogue, and fidelity to the full BST framework. These findings demonstrate the potential of an integrated AI system for child firearm safety training and outline key design considerations.**

***Index Terms*—Firearm Safety, Behavioral Skill Training (BST), AI-based Training.**

I. Introduction

Firearm-related injuries and deaths among children remain a critical public health concern in the United States [1]. Firearms are found in more than one-third of households, and roughly 74% of unintentional firearm incidents involving children are associated with firearms stored loaded, while 76% involve firearms that are stored unlocked [2], [3]. Approximately 5,000 children and adolescents are injured or killed by firearms each year, 56% of which occur in home environments and involve preventable access [3], [4]. When children discover firearms, research consistently shows that they are likely to handle or play with them [5], [6], highlighting the urgent need for effective firearm safety education.

Prior work has explored a range of approaches [5], [7]–[10]. While rule-based programs help children verbally recall

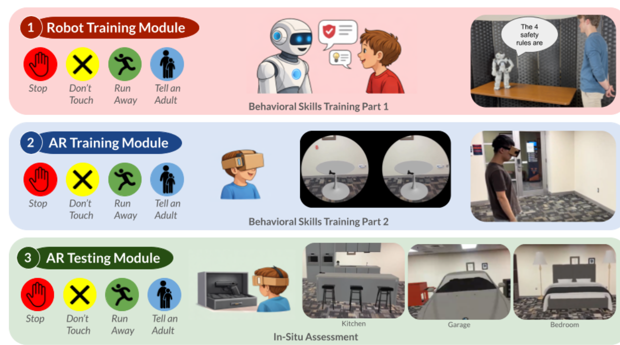

Fig. 1: Illustration of AI-Based firearm safety training system.

safety rules, children often fail to apply these behaviors in real-world situations [5]. More effective approaches use behavioral skills training (BST) and in-situ training (IST), which incorporate instruction, modeling, rehearsal, and feedback in realistic contexts [8], [10], [11]. BST provides structured, step-by-step skill acquisition through instruction, demonstration, practice, and corrective feedback, while IST extends this by embedding training within realistic environments to improve generalization and retention [6], [9]. However, both methods are resource-intensive, requiring trained facilitators and controlled environments, and do not consistently generalize across contexts or individuals [5], [6].

Advances in AI, augmented reality (AR), and social robotics present new opportunities to address these limitations. AR can simulate everyday environments such as homes or playgrounds, humanoid robots can deliver consistent and engaging instruction through expressive gestures and adaptive dialogue, and large language models (LLM)-based systems enable adaptive interactions grounded in expert knowledge [12]–[14]. Together, these technologies offer the potential to create immersive, scenario-based learning environments in which children can safely practice safety behaviors at scale, without the logistical demands of traditional training. However, limited research has examined how these technologies can be effectively integrated into firearm safety education for young children. In particular, prior work does not clearly address how AI, AR, and robotic systems can incorporate established behavioral training methods, how such systems are perceived by domain experts, or how expert knowledge can be embedded into their instructional design and delivery. Existing simulation-based efforts remain sparse, and no prior system unifies BST principles, AR-based in-situ rehearsal, and LLM-driven conversational instruction within a single platform for this age group. This gap motivates the need for a unified, interactive system that integrates evidence-based safety training with emerging technologies to improve both engagement and behavioral transfer. Our research is guided by the following research questions.

- **RQ1:** How can expert knowledge, existing curricula, and stakeholder requirements inform the design of an AI-based firearm safety education system for young learners?
- **RQ2:** How can AI, social robotics, and augmented reality be integrated to deliver behaviorally grounded firearm safety education?
- **RQ3:** How do domain experts evaluate the effectiveness and appropriateness of such a system, and what design implications emerge?

To answer these questions, we developed a NAO robot and AR-based training system that teaches children, aged 4-7 years old, the four-step firearm safety sequence (Stop, Don't Touch, Run Away, Tell a Trusted Adult) using LLM-driven interactive dialogue grounded in a retrieval-augmented knowledge base, along with an AR in-situ testing module across three domestic scenarios (e.g., kitchen, bedroom, and garage) (Fig. 1). We evaluated the system with ten experts in child development, education, and injury prevention. The system received favorable ratings across all modules (means: 4.25, 4.36, and 4.47), with strengths in usability, emotional safety, and scenario engagement. Our qualitative analysis identifies opportunities to improve the structural fidelity of BST, enhance developmental accessibility for younger children, and strengthen the ecological validity of the safety procedures. These results support the use of integrated robot-AR-LLM systems for evidence-based child safety education and point to critical design refinements needed for deployment.

II. Related Work

Firearm Safety Skills Training: Research on firearm safety education for young children has consistently demonstrated that active, practice-based approaches such as behavioral skills training (BST), a structured cycle of instruction, modeling, rehearsal, and feedback, outperform passive knowledge instruction [5], [7], [9], [11], [26]. Subsequent work extended these protocols to parent-implemented and simulation-based delivery, showing that active, practice-based approaches consistently outperform passive instruction. [6], [8].

However, the bulk of this foundational work dates to the early 2000s and continues to rely on human facilitators, while more recent efforts leveraging technology have explored passive alternatives that lack the real-time, adaptive interaction that defines effective BST, leaving a critical gap in leveraging modern technological advancements to reliably deliver the full BST cycle interactively and at scale.

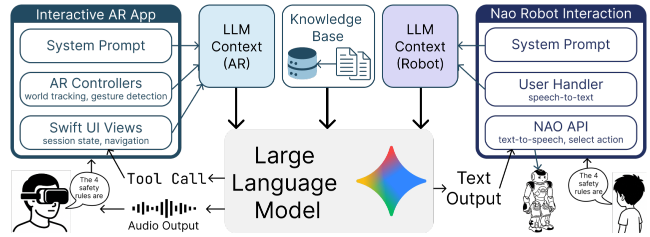

Fig. 2: System architecture of AR and robot-based training.

AR and VR-based Child Safety Training: Immersive simulation technologies, including augmented reality (AR) and virtual reality (VR), have been increasingly applied to safety education across a range of domains, demonstrating consistent benefits for engagement, skill acquisition, and generalization over traditional instruction [15], [17], [18]. Immersive simulation has been shown to be effective and produce skill generalization across settings, and has been identified as a viable approach for scaling BST delivery [18], with child-specific research further demonstrating acquisition and generalization of safety-critical behaviors across domains [16], [17]. While AR and VR show strong promise for child safety training, their application to firearm safety education for young children remains unexplored, representing a critical gap in scalable, immersive, and adaptive behavioral training.

Conversational AI and Robots for Learning: Social robots and LLM-powered conversational agents have been extensively studied as interactive learning partners for children, with research consistently demonstrating that adaptive, personalized interaction improves both engagement and learning outcomes [20]–[23]. Physical robot presence and social behavior have been shown to produce learning gains comparable to human tutoring on restricted tasks, largely due to engagement benefits that traditional learning technologies lack [20]. The integration of LLMs into social robots enables dynamic, open-ended instructional dialogue and adaptive interaction with children, with design requirements emphasizing age-calibrated cognitive scaffolding. [19], [22], [24].

While social robots and LLM-powered agents have demonstrated clear promise for child learning across general educational domains, a critical gap exists in deploying these systems for contexts where response accuracy is essential and deviations from evidence-based content have consequences.

III. Methodology

We developed a three-module firearm safety education system that integrates a humanoid NAO robot, augmented reality (AR), and large language model (LLM)-driven dialogue to deliver evidence-based training for children aged 4-7. The system follows Behavior Skills Training (BST) [7] across three modules: robot-delivered instruction, AR-based guided practice, and AR in-situ behavioral assessment, corresponding to the BST stages of instruction, rehearsal, and generalization. The system comprises four core components: domain knowledge acquisition, knowledge-grounded LLM dialogue, robot interaction design, and AR module design.

A. Domain Knowledge Acquisition

System development began with structured expert consultations involving firearm-safety experts, child psychologists, behavioral therapists, and public-health professionals to identify common real-world firearm exposure contexts and establish developmentally appropriate instructional norms. This input was combined with a review of established programs, Eddie Eagle GunSafe [27], the Virginia Finnegan Fox curriculum [28], and the McGruff crime-prevention series [29], which informed both the four-step safety sequence and age-appropriate dialogue design. Using the Finnegan Fox script as a structural template, we developed three parallel instructional scripts for the robot module, tailored to age subgroups within 4–7 by adjusting vocabulary complexity and question framing while maintaining a consistent pedagogical structure. Each script includes a contextual introduction situating firearms within familiar societal roles, followed by an interactive story and comprehension questions. All questions were designed to elicit open-ended responses, enabling assessment and reinforcement of conceptual understanding rather than simple recall.

B. Knowledge-Grounded LLM Dialogue

As shown in Figure 2, conversational intelligence across both the NAO robot and AR modules is powered by a shared large language model, selected for its balance of response speed and language quality in child-facing interaction contexts. To ground model responses in validated educational and clinical content rather than general knowledge, we constructed a shared retrieval-augmented generation (RAG) knowledge base consisting of 188 curated documents spanning firearm-safety literature, child developmental psychology, BST methodology, and existing educational program materials. The system combines the active discussion question with the child's transcribed speech to search the knowledge base for the most relevant passages, which are then provided to the model as reference material. This ensures that the system draws from the curated domain knowledge rather than general model knowledge.

C. Robot Interaction Design

The robot module operates as two coordinated processes: one running on a host computer that handles speech recognition, knowledge retrieval, and AI response generation, and another running on the robot that handles speech delivery and physical behavior. These processes communicate via a request–acknowledgment protocol, where the robot signals completion of each utterance before the next is sent. This split architecture accommodates compatibility constraints, as the NAO environment supports an older Python version while AI and speech libraries require a newer one. The robot is configured for posture initialization and live face tracking, with behavior scripted through custom code. Speech is delivered at a reduced rate with synchronized gestures corresponding to each safety step: a raised-hand stop ("Stop"), a negation motion ("Don't Touch"), a stepping-back movement ("Run Away"), and a pointing gesture ("Tell a Grown-Up") (Fig. 3). Pacing was refined through iterative testing with explicit pauses to match children's cognitive processing speeds.

D. AR Module Design

The AR system runs on a smartphone inserted into a stereoscopic headset, including low-cost cardboard options, supporting accessible real-world deployment. The application comprises three layers: a view layer for navigation and state-driven presentation, AR controllers for world tracking, object placement, and behavior detection, and voice/AI services for real-time interaction and session control. An AI voice coach listens and responds in real time, using structured session signals to advance, reset, or complete scenarios based on the child's actions while maintaining conversational flow. The Safety Training module follows a two-phase structure: a verbal recall phase in which the child recites the four safety steps, followed by AR-based practice once the coach confirms completion (Fig. 4). The Safety Testing module presents three sequential scenarios (kitchen, bedroom, and garage), each set in a friend's house to reflect high-risk real-world contexts. In each scenario, the child performs a task that leads to the discovery of an unsecured firearm. Correct responses receive immediate praise, while incorrect actions (e.g., reaching for the firearm) trigger corrective feedback and a scenario reset. Figure 4 shows a sample iteration in the AR training module.

IV. Study

A. Study Procedure

We conducted an expert evaluation in which adult professionals assessed each system module using a structured two-phase protocol aligned with the system's instructional sequence. Each session lasted approximately 60 minutes and was conducted remotely via Zoom. At the start of each session, participants provided informed consent, were assigned a unique study ID, and completed a demographics survey capturing professional background and domain expertise. Participants then moved through the three modules in sequence, evaluating the NAO robot module, then the AR training module, and finally the AR testing module. For each module, participants viewed a video walkthrough that demonstrated correct, incorrect, and unclear child interactions. Following each module, participants completed a short survey assessing dimensions relevant to that component, such as instructional clarity, tone, pacing, age-appropriateness, usability, scenario realism, and potential emotional impact on young children. Each session concluded with a brief semi-structured discussion allowing participants to elaborate on their survey responses and share additional observations.

B. Participants

We recruited ten expert participants ($n = 10$) with expertise in child development, education, healthcare, injury prevention, or firearm safety instruction. Participants were required to be at least 18 years old. We provide no compensation. The mean participant age was 42.8 years ($SD = 12.55$). Participants

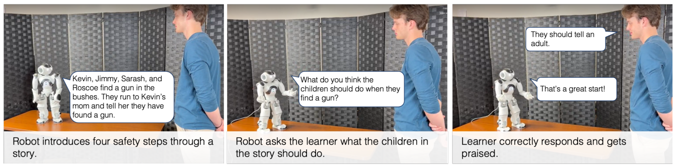

Fig. 3: Interaction in Robot-based training module.

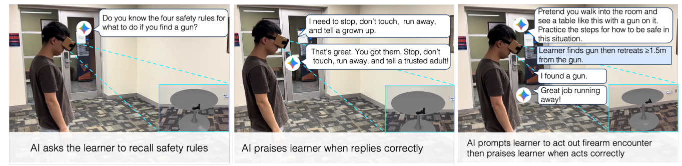

Fig. 4: Interaction in AR-based training module

self-reported expertise across four domains on a 1 (no experience) to 5 (expert-level) scale, with high experience in working with children ($M = 4.4$, $SD = 0.84$) and firearm safety or injury prevention ($M = 4.3$, $SD = 1.25$), and lower familiarity with robotics ($M = 1.8$, $SD = 0.63$) and augmented reality ($M = 2.1$, $SD = 0.99$).

C. Measures

We utilize a pre-task survey to collect information on professional background, domain expertise, and prior experience in social robotics, augmented reality, and child safety education. The post-module surveys assessed each system component using 5-point Likert-scale items (1 = Strongly Disagree, 5 = Strongly Agree) across five dimensions: instructional effectiveness, age-appropriateness, engagement and interaction quality, usability and effort, and emotional safety and ethics [25], [30], and included open-ended questions for qualitative feedback. For each module, item responses were averaged within each dimension to compute participant-level composite scores and category-level average ratings.

D. Implementation Details

The system was implemented using a NAO humanoid robot controlled via the NAOqi API and an AR application deployed on an iPhone with ARKit and LiDAR capabilities. Conversational interaction across modules was powered by Gemini, with AI-generated dialogue grounded in a retrieval-augmented generation (RAG) architecture using a curated knowledge base of 188 documents. Speech input was processed via Google's speech recognition API, and the robot delivered speech using the ALAnimatedSpeech API with synchronized gestures.

V. Results

A. Evaluation of Robot Module

The robot module received generally strong ratings across all five categories ($M = 4.25$, $SD = 0.35$) (Fig. 5 (a)). Usability and effort was rated highest ($M = 4.42$, $SD = 0.46$), indicating experts viewed the module as placing minimal demands on the child, followed closely by emotional safety and ethics ($M = 4.40$, $SD = 0.84$), affirming that the robot introduced safety content without inducing fear or distress. Engagement and interaction quality received positive ratings ($M = 4.30$, $SD = 0.59$), reflecting expert confidence in the robot's gestures, pacing, and conversational style for holding children's attention. Instructional effectiveness ($M = 4.12$, $SD = 0.66$) showed more variability, suggesting some disagreement on the module's standalone instructional impact. Age-appropriateness received the lowest rating ($M = 4.05$, $SD = 0.69$), with experts raising concerns about vocabulary complexity and pacing for children at the younger end of the target range, which are elaborated in the qualitative findings.

B. Evaluation of AR Module

The AR training module was evaluated across all five categories (overall $M = 4.37$, $SD = 0.43$) (Fig. 5 (b)). Emotional safety and ethics received the highest mean rating ($M = 4.65$, $SD = 0.41$), indicating consensus that the module's corrective feedback was proportional and appropriate. Usability and effort was also rated highly ($M = 4.55$, $SD = 0.33$) with the lowest variability category across this module, suggesting that experts viewed the gesture-based interaction as feasible for young children. Age-appropriateness ratings were favorable ($M = 4.25$, $SD = 0.68$). Engagement and interaction quality received positive but with more variable ratings ($M = 4.23$, $SD = 0.67$), with the variability reflecting mixed views on whether gesture-based practice alone would sustain children's attention over a full session. Instructional effectiveness was favorable ($M = 4.17$, $SD = 0.82$), though the variability indicates differing expert opinions on the module's adherence to the complete BST rehearsal cycle.

C. Evaluation of AR Testing Module

The AR simulation module received strong ratings across all five categories (overall $M = 4.46$, $SD = 0.38$) (Fig. 5 (c)). Engagement and interaction quality was rated highest ($M = 4.70$, $SD = 0.35$), reflecting expert confidence in the realistic scenario design and its capacity to immerse children in contexts that closely mirror real-world firearm encounters. Usability ($M = 4.60$, $SD = 0.66$) and emotional safety and ethics ($M = 4.45$, $SD = 0.51$) were also rated highly, with experts affirming the scenarios as feasible for controlled educational settings and appropriate in their handling of safety-critical content. Age-appropriateness ($M = 4.40$, $SD = 0.46$) and instructional effectiveness ($M = 4.30$, $SD = 0.60$) rounded out the ratings, with experts consistently agreeing that the scenarios could effectively guide children through the correct behavioral sequence and support skill generalization.

D. Qualitative Analysis

We conducted thematic analysis of open-ended survey responses and group discussions.

Language and Vocabulary Accessibility. Nearly every participant flagged at least one language issue, making this the most consistently raised concern across all three evaluation sessions. While participants praised the overall tone across all modules as appropriately safety-focused and non-punitive, the central tension was between terminology that is technically

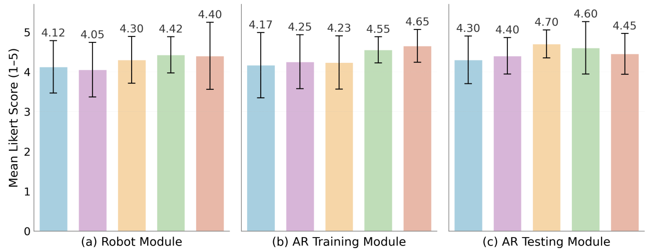

Fig. 5: Expert ratings across five dimensions for each module: Instructional Effectiveness (■), Age-Appropriateness (■), Engagement & Interaction (■), Usability & Effort (■), and Emotional Safety & Ethics (■).

accurate and language that is genuinely accessible to a 4-year-old. Participants noted "the term 'reinforce' is too advanced" and "'Local law enforcement authorities' could be too complex for kids," while the interchangeable use of 'gun' and 'firearm' risks confusion, with 'gun' considered more accessible.

Trauma-Informed Design. Concerns related to trauma-informed design emerged independently across multiple participants, who noted that some children in the target age range will have already been personally impacted by firearms, whether through community violence, family exposure, or prior traumatic experience. Participants asked "how the robot may respond to a child indicating how they or a loved one have been personally impacted." Experts recommended that a reactive protocol be established for cases where a child becomes distressed mid-session or discloses a traumatic experience to the AI, so that any disclosure of trauma would be met with acknowledgment before redirecting, rather than a return to the task as done with off-topic comments by a child.

AI Dialogue Quality and Conversational Robustness. Participants identified several specific failure modes in the AI's dialogue that could undermine the system's educational goals. The most concrete issue was the AI vaguely praising the child, which participants brought up as a concern because "a younger child would correlate the praise directly to finding a gun rather than running from it," and the system should "specifically say what they did and reinforce the four rules." Response length and pacing were also flagged as concerns, with participants noting that "kids . . . aged 3-5 don't have long attention spans".

BST Fidelity and Instructional Design. BST fidelity concerns were raised by participants with expertise in behavioral skills training, particularly that while the AR module supports rehearsal, the system omits the modeling phase. Experts also emphasized the need to include scenarios with a nearby adult to enable realistic "tell an adult" interactions, thereby ensuring that children perform actions as they would in real-world settings. Additional concerns focused on AR realism, including the need for more home-like environments and improved visual salience (e.g., low contrast between the firearm and background objects). Participants further critiqued the "what do you see" prompt, suggesting that placing the firearm near the search target without explicit cues would support more natural responses and create opportunities for implicit learning when the child ignores it.

VI. Discussion

Design from Expert Knowledge and Curricula (RQ1). The system's instructional content was grounded in established curricula including Eddie Eagle GunSafe, the Virginia Finnegan Fox, and the McGruff crime-prevention series, supplemented by consultation with firearm-safety experts, child psychologists, and public-health professionals. Expert evaluation affirmed this approach was largely successful, with strong emotional safety and ethics ratings across all modules and corrective feedback viewed as educational without being punitive (AI Dialogue Consistency). Vocabulary from existing curricula does not consistently translate to the youngest learners (Language Accessibility), with terms like "firearm" and "local law enforcement authorities" creating comprehension barriers for 4-year-olds, reflected in the lowest mean rating for the robot module. Experts also identified the need for trauma-informed protocols beyond existing curricula (Trauma-Informed Design), as some children will have prior personal exposure to firearms. Finally, the system does not yet implement the full BST cycle (BST Fidelity), with the modeling phase absent and the "tell an adult" step lacking a reportable recipient.

Integration of AI, Robotics, and AR (RQ2). The system maps directly to the BST stages of instruction, rehearsal, and generalization: the NAO robot delivers didactic instruction, the AR training module enables gesture-based rehearsal with real-time AI coaching via a RAG architecture, and the AR testing module presents realistic scenarios across three domestic settings. Quantitative results support the viability of this integration, with consistently high usability and effort ratings across all modules, strong engagement ratings for the AR simulation module, and experts affirming that RAG-constrained dialogue was generally appropriate and on-topic. The AI dialogue exhibited failure modes, praising the discovery event than the safety behavior and inconsistent step ordering (AI Dialogue Consistency), suggesting that an output validation layer is needed to enforce behavioral consistency at the instructional level.

Expert Evaluation and Design Implications (RQ3). Expert evaluation produced a consistent pattern across both quantitative and qualitative data: the system's instructional design and emotional safety were viewed favorably, while concerns centered on developmental accessibility for the youngest learners and specific gaps in behavioral training fidelity. Quantitatively, category-level Likert ratings were high across all five evaluation categories for every module. Emotional safety and ethics received consistently strong ratings, confirming that experts viewed the system as introducing safety-critical content appropriately. Instructional effectiveness showed the most variability across modules, reflecting differing expert opinions on the system's adherence to the complete BST cycle and on whether individual modules could achieve their instructional goals independently.

Design Implications. Our analysis highlights key gaps in language accessibility, trauma-informed design, AI dialogue consistency, and BST fidelity, informing the following design implications. First, the system should adopt age-stratified language and content profiles, tailoring vocabulary, sentence structure, and content complexity for younger (4–5) and older (6–7) children. Second, it should incorporate a proactive trauma-informed response layer that detects and appropriately responds to personal disclosures, including mechanisms for facilitator intervention. Third, instructional consistency should be enforced through a structured validation layer that ensures correct step ordering, parity, and feedback aligned with safety behaviors. Fourth, the system should complete the BST cycle by incorporating a modeling phase and enabling a realistic "tell an adult" interaction through embedded characters or agents.

VII. Conclusion

In this paper, we presented an integrated firearm safety education system combining a NAO humanoid robot, augmented reality, and LLM-driven dialogue to deliver evidence-based behavioral skills training to children. The system comprises three BST-aligned modules: robot-delivered instruction, AR-guided rehearsal, and AR in-situ simulation. We evaluated the system through expert review with ten professionals, and qualitative analysis identified four key areas for refinement: language accessibility, trauma-informed design, AI dialogue consistency, and BST structural fidelity. Limitations include the small and demographically narrow expert sample, lack of iterative design cycles, and absence of testing with child participants, leaving real-world efficacy an open question. These findings provide expert-validated evidence of the system's viability while highlighting necessary refinements prior to child deployment. Future work will focus on child-participant studies, implementing identified improvements, and exploring lower-cost hardware for broader deployment.

References

[1] Centers for Disease Control and Prevention, National Center for Injury Prevention and Control, "Fast Facts: Firearm Injury and Death," *Centers for Disease Control and Prevention*, 2024. Available: https://www.cdc.gov/firearm-violence/data-research/facts-stats/index.html. Accessed: Apr. 14, 2026.

[2] R. Miltenberger, M. A. Novotny, T. Maxfield, and R. Baruni, "Evaluation of video self-modeling to teach firearm safety skills," *Behavioral Interventions*, vol. 37, no. 3, 2022.

[3] R. F. Wilson, S. Mintz, J. M. Blair, C. J. Betz, A. Collier, and K. A. Fowler, "Unintentional firearm injury deaths among children and adolescents aged 0–17 years — National Violent Death Reporting System, United States, 2003–2021," *MMWR Morbidity and Mortality Weekly Report*, vol. 72, no. 50, 2023.

[4] B. J. Naik-Mathuria, C. M. Cain, E. A. Alore, L. Chen, and L. A. Pompeii, "Defining the full spectrum of pediatric firearm injury and death in the United States: It is even worse than we think," *Annals of Surgery*, vol. 278, no. 1, 2023.

[5] M. B. Himle, R. G. Miltenberger, B. J. Flessner, and B. Gatheridge, "Teaching safety skills to children to prevent gun play," *Journal of Applied Behavior Analysis*, vol. 37, no. 1, 2004.

[6] M. E. Orner, R. G. Miltenberger, and T. Maxfield, "Evaluating small-scale simulation training of firearm safety to children with autism spectrum disorder," *Behavioral Interventions*, vol. 36, no. 4, 2021.

[7] R. G. Miltenberger, "Teaching safety skills to children: Prevention of firearm injury as an exemplar of best practice in assessment, training, and generalization of safety skills," *Behavior Analysis in Practice*, 2008.

[8] M. A. Novotny, R. G. Miltenberger, R. R. Baruni, T. C. Maxfield, V. Larson, and S. Yampolskaya, "Evaluating parent-implemented behavioral skills training for teaching firearm safety skills," *Behavioral Interventions*, vol. 38, no. 1, 2023.

[9] R. G. Miltenberger, C. Flessner, B. Gatheridge, B. Johnson, M. Satterlund, and K. Egemo, "Evaluation of behavioral skills training to prevent gun play in children," *JABA*, vol. 37, no. 4, 2004.

[10] R. G. Miltenberger, B. J. Gatheridge, M. Satterlund, K. R. Egemo-Helm, B. M. Johnson, C. Jostad, P. Kelso, and C. A. Flessner, "Teaching safety skills to children to prevent gun play: An evaluation of in situ training," *JABA*, vol. 38, no. 3, 2005.

[11] B. J. Gatheridge, R. G. Miltenberger, D. F. Huneke, M. J. Satterlund, A. R. Mattern, B. M. Johnson, and C. A. Flessner, "Comparison of two programs to teach firearm injury prevention skills to 6-and 7-year-old children," *Pediatrics*, vol. 114, no. 3, 2004.

[12] E. Dempsey et al., "Immersive virtual reality for unintentional injury prevention training with children: A systematic review," *ScienceDirect*, 2024.

[13] Z. Lan et al., "Large Language Models for Education: A Survey and Outlook," *arXiv preprint arXiv:2403.18105*, 2024.

[14] S. Naidu, E. Smith, C. Hagood, A. Rolly, S. Sarker, C. Hayes, and T. Iqbal, "A Data Capture and Gesture Recognition System to Enable Human-Robot Collaboration," in *SIEDS*, 2025.

[15] P. Gong, Y. Lu, R. Lovreglio, X. Lv, and Z. Chi, "Applications and effectiveness of augmented reality in safety training: A systematic literature review and meta-analysis," *Safety Science*, vol. 178, 2024.

[16] D. C. Schwebel, A. Johnston, D. McDaniel, and L. A. McClure, "Child pedestrian safety training in virtual reality: How quickly do children achieve adult functioning and what individual differences impact learning efficiency?" *Journal of Safety Research*, vol. 88, 2024.

[17] P. Araiza-Alba, T. Keane, and others, "The potential of 360-degree virtual reality videos to teach water-safety skills to children," *Computers & Education*, vol. 163, p. 104096, 2021.

[18] A. A. Chang, E. Kazemi, V. Esmaeili, and M. S. Davies, "The effectiveness of virtual reality training: A systematic review," *Journal of Organizational Behavior Management*, vol. 44, no. 3, 2024.

[19] N. Kurian, "Developmentally aligned AI: A framework for translating the science of child development into AI design," *AI, Brain and Child*, vol. 2, no. 1, 2025.

[20] M. I. Ahmad, O. Mubin, S. Shahid, and J. Orlando, "Robot's adaptive emotional feedback sustains children's social engagement and promotes their vocabulary learning: A long-term child-robot interaction study," *Adaptive Behavior*, vol. 27, no. 6, 2019.

[21] T. Belpaeme, J. Kennedy, A. Ramachandran, B. Scassellati, and F. Tanaka, "Social robots for education: A review," *Science Robotics*, vol. 3, no. 21, p. eaat5954, 2018.

[22] M. Elgarf, H. Salam, and C. Peters, "Fostering children's creativity through LLM-driven storytelling with a social robot," *Frontiers in Robotics and AI*, vol. 11, p. 1457429, 2024.

[23] Q. Jiang, L. Zak, S. Leshem, P. Rampa, S. Howle, H. N. Green, and T. Iqbal, "Embodied AI for Financial Literacy Social Robots," in *SIEDS*, 2023.

[24] C. Miu, J. Gopurathingal, V. Thota, M. Thompson, N. van Beek, J. Kuczynski, J. Gadewadikar, and T. Iqbal, "A Financial Literacy AI-Enabled Voice Assistant System for Educational Use," in *SIEDS*, 2022.

[25] R. R. Baruni and R. G. Miltenberger, "Teaching safety skills to children: A discussion of critical features and practice recommendations," *Behavior Analysis in Practice*, vol. 15, 2022.

[26] R. Miltenberger, A. Gross, P. Knudson, A. Bosch, C. Jostad, and C. Breitwieser, "Evaluating behavioral skills training with and without simulated in situ training for teaching safety skills to children," *Education and Treatment of Children*, vol. 32, 2009.

[27] Eddie Eagle GunSafe Program. *Program Resources*. National Rifle Association. Available: https://eddieeagle.nra.org/program-resources/. Accessed: Apr. 10, 2026.

[28] Elementary School Gun Safety Guidelines and Curriculum. *Virginia Department of Education*. Available: https://www.doe.virginia.gov/programs-services/school-operations-support-services/safety-crisis-management/gun-safety. Accessed: Apr. 10, 2026.

[29] McGruff Talks to Students About Gun Safety. *Young Minds Inspired*. Available: https://ymiclassroom.com/lesson-plans/mcgruff-gun-safety/. Accessed: Apr. 10, 2026.

[30] Developmentally Appropriate Practice (DAP) Position Statement. *National Association for the Education of Young Children*. Available: https://www.naeyc.org/resources/position-statements/dap/contents. Accessed: Apr. 10, 2026.

[31] V. Braun and V. Clarke, "Using thematic analysis in psychology," *Qualitative Research in Psychology*, vol. 3, no. 2, 2006.