"If you are a Star Wars fan, use the force": Exploring Children's Virtual-Object Interaction Preferences in AR Headsets

Alejandro Delgado
College of AI, Cybersecurity and Computing
University of South Florida
Tampa, Florida, USA
alejandrodel@usf.edu

Md Mehedi Hasan Jibon
College of AI, Cybersecurity and Computing
University of South Florida
Tampa, Florida, USA
jibon@usf.edu

Hetvi Shah
College of AI, Cybersecurity and Computing
University of South Florida
Tampa, Florida, USA
hetvi@usf.edu

Fareeza Rahman
College of AI, Cybersecurity and Computing
University of South Florida
Tampa, Florida, USA
fareeza@usf.edu

Anzhelika Kurnikova
College of AI, Cybersecurity and Computing
University of South Florida
Tampa, Florida, USA
akurnikova@usf.edu

Julia Woodward
College of AI, Cybersecurity and Computing
University of South Florida
Tampa, Florida, USA
juliaevewoodward@usf.edu

Abstract

Children are increasingly using augmented reality (AR) headsets in different contexts, such as education. However, it is unclear how children expect to interact with virtual objects in AR headsets; children's expectations for technology can significantly differ from adults. Therefore, we conducted an elicitation study with 20 children (ages 9-12), in which children proposed interactions for tasks with a virtual cube (e.g., moving, expanding, creating, etc.) in an AR headset. We constructed a conceptual model of children's expectations with virtual-object interactions in AR headsets and analyzed their proposed interactions. We found that children preferred gestures, expecting to utilize their whole body (e.g., pushing, kicking) and external objects (e.g., hammer, sword) to interact with the cube, and rarely considered speech, which differs from adults. Children also frequently added their own motivations, creating a narrative behind their interactions. We provide foundational insights into children's expectations for virtual-object interaction in AR headsets.

CCS Concepts

• **Human-centered computing → Mixed / augmented reality**; *HCI theory, concepts and models.*

Keywords

Children, augmented reality, augmented reality headsets, conceptual model, gesture elicitation

ACM Reference Format:
Alejandro Delgado, Md Mehedi Hasan Jibon, Hetvi Shah, Fareeza Rahman, Anzhelika Kurnikova, and Julia Woodward. 2026. "If you are a Star Wars fan, use the force": Exploring Children's Virtual-Object Interaction Preferences in AR Headsets. In *Proceedings of the 2026 CHI Conference on Human Factors in Computing Systems (CHI '26), April 13–17, 2026, Barcelona, Spain.* ACM, New York, NY, USA, 16 pages. https://doi.org/10.1145/3772318.3790552

1 Introduction

Augmented reality (AR) is a technology that supplements the real world by combining virtual objects with the natural environment [70]. Compared to other AR platforms (e.g., smartphones, tablets), AR headsets offer increased user immersion and hands-free capabilities, making them a popular choice [2, 49, 72]. Children are increasingly using AR headsets in different contexts, such as for educational tools [28, 33, 35, 65], to help support children with autism and ADHD [6, 24, 31, 62], to reduce stress during medical procedures [11], and to improve gait rehabilitation [23]. Evidence suggests that AR headsets may enhance children's learning, engagement, and motivation [28, 33, 46]. While children are benefiting from using AR headsets, they can still experience difficulty interacting with the headset and virtual content [3, 28, 57].

Prior work has started to investigate how adults expect to interact with virtual objects in AR headsets [16, 53, 55, 78]. However, there is evidence that children's interaction behaviors and expectations are different from adults' across multiple technologies, such as touchscreen devices [4, 61, 87], spherical displays [66], voice assistants [36, 37], and AR headsets [86], which can lead to errors and children's frustration when they interact with these technologies. It is unclear how children expect to interact with virtual objects in AR headsets. Understanding children's expectations will allow designers to design more intuitive AR headset applications. According to the Expectation-Confirmation Model for information systems, users experience more satisfaction with a system when they perceive it as useful and their expectations are met [8].

To better understand children's natural expectations and preferences for virtual-object interactions in AR headsets, we conducted an open-ended AR headset elicitation study with 20 children (ages 9-12). Gesture elicitation has been shown to be an effective methodology for understanding children's interaction expectations with a variety of technologies, including whole-body interfaces [14], touchscreen devices [61], and spherical displays [66]. To our knowledge,

CHI '26, Barcelona, Spain

ACM ISBN 979-8-4007-2278-3/26/04
https://doi.org/10.1145/3772318.3790552

we are the first to explore children's expectations of virtual-object interactions through the use of gesture elicitation with an AR headset. We contribute foundational empirical insights into children's interaction preferences, extending prior conceptual work by studying children's interactions with an AR headset. In our elicitation study, children saw a virtual cube in the AR headset perform 17 different effects (e.g., moving, expanding, creating, etc.) and we asked them to provide an interaction they believed would replicate the effect they saw. We opted for a cube because its simple geometry supports a clear visual interpretation of transformations such as translation, scaling, and rotation, and it allows us to compare our findings to prior AR headset elicitation studies with adults that utilized a virtual cube [16, 79]. After the children proposed an interaction, we asked the children open-ended questions about their interactions to better understand their reasoning and preferences.

After the study was completed, we analyzed the children's explanations through affinity diagramming, a bottom-up inductive approach used to analyze large-scale qualitative data [7]. Based on our analysis, we constructed a conceptual model of children's expectations with virtual-object interactions. Additionally, we performed qualitative coding on the children's interactions, allowing us to provide a quantitative analysis of children's interaction preferences to support our conceptual model. We found that a majority of the children preferred gestures, interacting with the cube as if it were a real object (i.e., direct manipulation), as well as utilizing their whole body (e.g., pushing, kicking) and external objects (e.g., microscope, hammer, ninja stars, screwdriver, sword) to interact with the cube. Furthermore, the children rarely used speech in their interactions, which differs from what has been found with adults [16]. Our conceptual model outlines how children's internal factors and expectations, as well as the technology and environment, influence the children's interaction choices. We also found that children created a narrative behind their interactions, contextualizing their actions and attributing a purpose based on their own intrinsic motivations. The contributions of our work include: (1) To our knowledge, we are the first to explore children's expectations of virtual-object interactions through the use of gesture elicitation with an AR headset, (2) a conceptual model outlining children's perceptions and reasoning behind their AR headset virtual-object interactions, and (3) a set of design recommendations to improve future AR headset experiences for children. Our work contributes foundational empirical insights into children's AR headset interaction preferences, which can aid designers and researchers in creating AR headset applications for children.

2 Related Work

We focus our related work on three categories: (1) exploring the use of elicitation methodologies to improve technology and better understand users, (2) examining prior work on children using augmented reality (AR) headsets, and (3) the ethics of using AR headsets with children.

2.1 Gesture Elicitation Studies

Gesture elicitation is a methodology in which participants are shown an *effect* (i.e., referent) and are prompted to elicit a gesture they believe will replicate the observed effect [80]. For example, in Wobbrock et al.'s [81] gesture elicitation study with touchscreen devices, the participants were shown an animation of a field of objects moving from left to right (i.e., referent) and then asked to perform a gesture they believe would replicate the animation they saw. Elicitation studies allow researchers to determine a set of gestures that are easier to learn, perform, and recall for users [80, 81]. Prior work has utilized gesture elicitation to improve various technologies such as mobile devices [18, 19, 45, 47, 60, 64], tabletop computers [81], AR windshield displays [12], whole-body interfaces [48, 56] and AR headsets [16, 32, 53, 78]. Although the aforementioned AR headset studies used gesture elicitation to study people's interaction expectations, they only focused on adults.

Researchers have started utilizing gesture elicitation to improve children's interactions with technology. Connell et al. [14] conducted an elicitation study with a whole-body interface in which children (ages 3 to 8) were shown a virtual cube on a TV screen completing different tasks, such as the cube getting bigger, rotating, etc. The researchers found that instead of using gestures to directly manipulate the object on screen, the children often simulated the object's movement path with their own bodies. Rust et al. [61] utilized gesture elicitation to better understand child-defined gestures and children's mental models for touchscreen tabletop interactions, in which participants (both adults and children) completed a set of tasks (move, pick many, pick one, remove, etc.). The authors found that while children were able to successfully elicit gestures, they used less symbolic gestures (e.g., drawing an undo arrow for remove) compared to adults. Soni et al. [66] explored the possibility of generalizing user-defined gestures used in flatscreens to interactive spherical displays for both adults and children (ages 7 to 11). By using gesture elicitation, the authors determined that flatscreen gestures could not be generalized to spherical displays. The authors also found that children were more likely than adults to use their hands interchangeably, propose gestures with dynamic hand poses (i.e., the hand pose changes as the hand moves), and favor continuous gestures over discrete ones (i.e., the effect of the gesture takes place while the gesture is being performed rather than after the gesture is completed).

These studies confirm that gesture elicitation can be an effective method for understanding children's unique mental models and expectations across various technologies. However, this user-centered approach has not yet been applied to determine how children expect to interact with AR headsets, in which they are not manipulating a screen but rather interacting with virtual objects that are overlaid on the real world.

2.2 Children and AR Headsets

AR headsets are being used as educational tools for children (e.g., games [3, 28], virtual field trips [65], museum exhibits [35]), as communication and emotion recognition aids for children with autism [6, 24, 62], as a way to help children with ADHD increase attention rates [31], and to help children relax during medical procedures [11]. For example, Juan et al. [28] created an AR headset-based game for children (ages 7 to 12) to learn about endangered animals, and Andersen et al. [3] designed the BattleBoard 3D AR board game with children to study design issues for AR tabletop play. Lauer et al. [33] examined using AR headsets with elementary school children and found that it had positive effects on their activity-related achievement emotions. Kim et al. [31] found that an eye-contact game on an AR headset could help children (ages 8-10) with ADHD increase their attention span.

Prior work has started to explore children's mental models [26, 82] and interactions [43, 44, 86] with AR headsets. Woodward et al. [82] conducted online participatory design (PD) sessions with children (ages 7-12) to explore their perceptions of using AR headsets for different tasks (e.g., chores, homework). The findings revealed that children perceive AR headsets as highly intelligent systems capable of recognizing and transforming their surroundings, providing an immersive experience. However, since the study was online, children did not interact with commercial AR headsets. Hourcade et al. [26] conducted play-based design sessions with preschool children (ages 3–5) in which children tried on low-fidelity AR smart-glasses prototypes and discussed what they would want the glasses to do. The authors found that children envisioned AR wearables helping with pretend play (e.g., imagining animal companions) and simple everyday activities (e.g., identifying objects or showing where to go). Bülbül and Özdinç [10] studied preschool children's experiences with tablet-based mobile AR activities in which animated 3D characters appeared when children scanned AR cards. They found that while children enjoyed the animated AR content, they still preferred real objects for tasks involving physical manipulation or sensory engagement. In terms of children's interactions, Munsinger and Quarles [43] examined interaction methods (i.e., voice, gesture, controller) for a Fitts' Law task in an AR headset with children (ages 9-11). The researchers discovered that controller selection was faster than both voice and gesture. Although the authors compared existing AR headset interaction methods with children, they did not investigate how children expect to interact with virtual objects in AR headsets. Woodward and Ruiz [86] examined how different textual designs in an AR headset affect adults' and children's (ages 9 to 12) task performance. The authors found significant effects of the design of information on children's task performance, such as a significant effect of information location on information recall accuracy. Specifically, when the information was placed in the main direction the children were looking at, it led to higher information recall accuracy. The authors did not find any significant effect of the design of textual information on adults' task performance, illustrating that information in AR headsets may have to be designed differently for children depending on the context.

Although prior work has started investigating AR headsets with children, there is still a gap in our understanding as to how children expect to interact with virtual objects while wearing an AR headset and why they have those expectations. For instance, while Woodward et al. examined children's perceptions of AR headsets, the children did not view or interact with any AR headset [82]. Moreover, Hourcade et al. and Bülbül and Özdinç [10, 26] focused on preschoolers' conceptual models and preferences for AR smart glasses and AR activities rather than gesture-level interaction with an AR headset. It is important to investigate how children may require different designs to meet their specific expectations, conceptual models, and needs. To our knowledge, we are the first to explore children's expectations of virtual-object interactions through the use of gesture elicitation with an AR headset.

2.3 Ethics of Using AR Headsets with Children

While virtual reality (VR) and augmented reality (AR) headset makers have traditionally recommended their devices for ages 13 and older due to a lack of research and consideration for younger users, children are still using these technologies (as mentioned above). Also, Meta has recognized that research is needed to provide age-appropriate immersive headset experiences for children and has started to incorporate parent-controlled accounts for children under 13 [40], as children are utilizing these technologies. For instance, VR is being widely used with children, as young as 4 and 5 years old, in medical contexts to help with pain and anxiety [13, 50, 63, 71] and for education [51, 76]. AR headsets are also increasingly being used by children as young as 4 years old [62], and have the potential to increase children's engagement, motivation, and learning [28, 33]. While VR headsets have been reported to induce negative effects such as cybersickness, even with children [29, 69], AR headsets have been reported to induce less symptoms since users are still grounded in the real-world [33, 62, 77, 88]. For example, Sahin et al. [62] did not find any negative effects of AR headsets on adults and children (as young as 4 years old) with autism. AR headsets — specifically, optical see-through headsets — are also more beneficial for safety for children, when compared to VR headsets, as children are still completely aware of their real-world surroundings. Prior work has identified ethical concerns for VR headsets, such as sensor vulnerability, social isolation, and desensitization as VR takes users outside of the real-world [21]. However, these ethical concerns are minimized in AR headsets, as users are still grounded and present in the real-world. Prior work found that AR headsets comforted children during a medical procedure, more so than a VR headset, as the children could still see their parents [11]. Southgate et al. [67] explored the ethics for designing AR experiences for children. The authors remarked that there has been little research on the physical, cognitive, and emotional effects of highly immersive experiences on children. Southgate et al. [67] emphasized considering the developmental stage of the children when designing and including researchers with expertise with children. For our study, we follow Southgate et al.'s ethical guidelines [67], such as obtaining IRB approval, parental informed consent, child assent, enforcing breaks from the headset, and allowing children to take breaks and stop whenever they would like. We further discuss our study procedure and child assent in the Methodology section below. Future research should focus on the effects of highly immersive experiences on children and how to create child-centered AR headset experiences.

3 Methodology

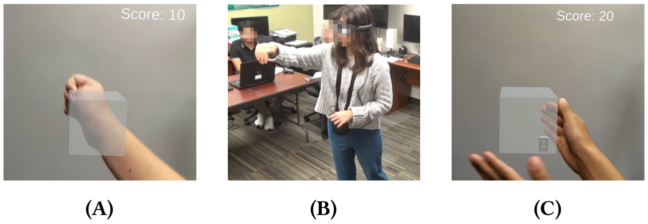
We conducted a within-subjects open-ended AR headset elicitation study with children to explore how they expect to interact with virtual objects in AR. In our study, 20 children observed a virtual cube perform 17 different effects (i.e., referents), such as the cube moving left, in an AR headset (Fig. 1). We chose a cube because its simple geometry supports a clear visual interpretation of transformations (e.g., rotation) without biasing participants toward specific interactions, and it allows us to compare our findings to prior AR headset elicitation studies with adults that used a virtual cube [16, 79].

To select the referents, we analyzed 42 existing gesture elicitation studies (e.g., [5, 14, 16, 18, 45, 53, 66, 78, 79, 81, 89]). We extracted 145 tasks and plotted them into a frequency matrix to identify the most common referents used in prior gesture elicitation studies. Finally, we selected the 17 most common referents that could be applied to a virtual cube in an AR headset (Table 1). We grouped these referents into three categories: (1) *Translation*, for movements of the cube; (2) *Manipulation*, for changes to the cube's properties (e.g., scale or rotation); and (3) *Abstract*, for higher-level or symbolic actions. This three-part categorization was informed by prior work on interaction taxonomies in 3D and AR contexts [16, 78]. We counterbalanced the order of the three referent groups between children to reduce ordering effects, while keeping the order of referents within each group fixed. For each referent, children first proposed an interaction to replicate what they saw (Fig. 1). Similar to prior work [66], we prompted children to provide a second interaction choice for the same referent, allowing them to skip one or both choices if they could not think of an interaction for that referent.

We conducted all studies in a spacious research lab, with at least two researchers present at all times. Each study was audio and video recorded. Our research protocol was approved by our Institutional Review Board, and we collected both parental consent and child assent before participation. Both the parents and children were informed of the risks of AR headsets. For instance, during the child assent process, we explicitly informed each child that "You may become tired or start to not feel well, such as getting a headache or nauseous. If that is the case let us know if you want a break or if you want to stop at any time." Each study lasted around 60 minutes, including the time taken during the consent process, break times, and debrief. Each study was split into three shorter sessions (i.e., one session for each referent category), with enforced breaks between each section in which we had the children remove the AR headset. Each child wore the AR headset for an average of 11 minutes (m) and 57 seconds (s) (SD=2m:32s) before taking it off for a break. The children were compensated with $40 and a small prize (e.g., bouncy balls, stickers).

Table 1: List of the 17 referents used in our study.

Group	Referents
Translation	Move (Up/Down) Move (Left/Right) Move (Towards/Away)
Manipulation	Make (Bigger/Smaller) Rotate (X/Y/Z) axis Zoom (In/Out)
Abstract	Create Destroy Select (One/Multiple)



(A) (B) (C)

Figure 1: Examples of the children's interactions: (A) punching the cube and destroying it; (B) jabbing the cube to select it; (C) grabbing and rotating the cube.

3.1 Participants

We recruited 20 children (16 male, 4 female) between the ages of 9 and 12 (M = 10.55, SD = 1.16), which is consistent with prior work [14, 43, 61, 66]. The children's grade levels spanned from 3rd grade to 7th grade. We used an email listserv, posted flyers on community social media sites, shared study information through a local coding program, and relied on word of mouth to reach families. Parents reached out to the researchers by email to schedule their child's study. Of the 20 children, 17 self-reported as right-handed. In addition, 17 of the children self-reported prior experience using a virtual reality (VR) headset, mostly for playing games such as *Beat Saber* and *Gorilla Tag*. Two children had previously used an AR headset: one during a prior research study and another while playing a game.

3.2 Study Design

The study used a within-subjects design, meaning each child completed all referents. We first verbally asked the children demographic questions, such as their age, grade level, and prior experience with VR and AR headsets. After the questions, the children then put on the AR headset and completed a short practice in which they saw the virtual cube in the headset and could walk around the room. We did the practice so the children could become comfortable with the AR headset and to reduce the novelty effect. During practice time, we verified with each child that the headset fitted them comfortably, making sure the head strap was properly adjusted and no cables were in the way. Additionally, we set the headset's viewing mode to "transparency", which is optimized for real-world use and increased viewing comfort. We also informed children that they were able to request as many breaks as they wanted during the study and could stop participation at any time.

The virtual cube appeared in front of the child's eyeline, aligned with their height (Fig. 1). We asked the children to stand on an "X" marked on the floor while the referent animation was being played to make sure the children always viewed the referent from the same angle and position. We explicitly mentioned to the children that they were free to move around after the animation played to perform their interaction. To support unconstrained responses, we informed the children at the beginning of the study that they could interact with the cube in any way they desired to replicate what they saw. We did not mention any specific capabilities of the AR headset as we did not want children's interaction proposals to be biased by the current technological limitations of the headset. For each referent, we prompted the children by saying: *"Pretend you are making the cube [Referent, e.g., move right]."* The headset then showed the animation of the cube performing that referent. After the animation, we asked: *"How would you make the cube [Referent] like you saw?"* The children then showed us how they would complete the interaction (e.g., physically acting out the interaction, giving a verbal command). We then followed up with: *"Why did you want to [Referent] the cube like that?"* We used open-ended questions in order to capture the children's natural interactions and reasoning. After the children proposed their interaction and explained their thought process, we asked them to rate their interaction proposal from 1 to 5 based on *Goodness of fit* ("The interaction I picked is a good match for its intended use") and *Ease of use* ("The interaction I picked is easy to do"), in alignment with prior research [16, 66, 81]. For ratings, we used a 5-point Likert scale with smiley faces, as in prior work with children [59]. However, Hall et al. [25] found that even with smiley face ratings, children are less likely to choose negative or neutral options; therefore, we opted for the happiness rating scale [25], which achieves more insight and variance with children. In the happiness rating scale, a rating of 1 shows a smiley face and a rating of 5 shows a very happy smiley face. Next, we asked for a second interaction choice for that referent: *"How would you make the cube [Referent] in another way?"* We then repeated the process: the interaction, explanation, and ratings. After the completion of each referent group (i.e., Translation, Manipulation, Abstract), we enforced a break in which the children removed the AR headset.

3.3 Equipment and Software

We used the Magic Leap 2 AR headset [38], as it is lightweight and designed for extended use. The Magic Leap 2 is lighter (i.e., 260 grams) [34] than other commercial AR headsets (e.g., HoloLens 2 is 566 grams) [41], which is more beneficial for children. We chose an optical see-through AR headset, as it does not block the children's periphery (important for safety) and feels similar to looking through glasses. The Magic Leap 2 has a field of view of 45H x 55V (70D) and a resolution of 1440 × 1760 pixels per eye. We built our AR app in Unity (version 2022.3.22f1) using the Magic Leap 2 SDK [38]. The app spawned a virtual cube with 10-centimeter sides, positioned 60 centimeters in front of the headset. All animations kept the cube within the device's field of view to avoid visual clipping. The app also included a simple scoreboard for gamification; incorporating gamification elements in empirical studies has been shown to increase study completion rates for children [9]. The Magic Leap headset also recorded the children's point-of-view video during each interaction. Details about the application and the referent animations are included in the Supplementary Materials.

3.4 Data Analysis

For analysis, we first qualitatively examined the children's mental models for how they expect to interact with a virtual object in AR headsets, and then quantitatively analyzed the children's proposed interactions. Our data included video and audio recordings of each child, along with point-of-view footage captured from the AR headset.

3.4.1 Qualitative Analysis. We analyzed children's verbal explanations to understand their mental models of interacting with virtual objects in AR headsets. The first two authors transcribed children's verbal explanations by reviewing the video and audio recordings. Similar to prior work [16, 82, 84], we analyzed the explanations through affinity diagramming, which is a method to organize large-scale qualitative data through a bottom-up inductive approach [7]. We did not compute inter-rater reliability as it is not appropriate or necessary when the research goal is to generate themes inductively rather than to test agreement [39]. To create the affinity diagram, the first two authors and the last author iteratively grouped the individual utterances into themes over the course of nine meetings (approximately 22 hours) using Lucidchart, a collaborative online whiteboard [1]. In total, we analyzed 655 explanations (20 children × 17 referents × 2 interactions per referent); 25 explanations were omitted due to missing or incomplete data. These omissions also included instances in which the child was unable to think of an interaction or explanation. With our affinity diagram analysis, we uncovered recurring themes and patterns in how children conceptualize interaction with virtual objects in AR headsets. We grouped the data into 9 overarching themes, which we linked to create our conceptual model (Fig. 2); described in the Results section.

3.4.2 Quantitative Analysis. We analyzed the children's video-recorded proposed interactions using qualitative coding, a method for capturing interaction properties and organizing them into codes [74, 83]. The creation of our initial codebook and selection of interaction modalities was based by prior work on user-driven elicitation [58], multimodal interaction [78], and gesture elicitation in AR and VR contexts [16, 17, 30, 53]. Additionally, we iteratively updated our codebook based on repeating patterns and important behaviors we identified during our study sessions. For each interaction, we identified the interaction modality (e.g., gesture, speech) and a corresponding action (e.g., push, pull, single word voice command) from that modality. We coded children's interactions across six modalities: (1) *Gesture* interactions were defined as any movement of the arms and/or hands with the intention to complete the referent. We coded the type of action (e.g., push, grab, tap) and any use of external objects (e.g., baseball bat, microscope); (2) *Speech* interactions were defined as words/commands that come from the child to complete the referent. We coded the verbatim utterance and the speech command type (single word, multi word, complete sentence); (3) *Body* interactions, in which the child's body physically moved to another position (e.g., running) or used their whole-body for the interaction (e.g., kicking, elbowing). We coded the specific physical movement; (4) *Eye* interactions were coded when the child specifically mentioned that they will use eye movement to complete the referent. We coded the specific behavior (e.g., blink, gaze, eye movement direction); (5) *Head* interactions were coded when the child specifically mentioned that they use their head to complete the referent. We coded the head's movement direction; and (6) *Other* for when the interaction did not fit any aforementioned modality. Our coding process also supported multimodal interactions, such as combining gestures with speech.

In total, we reviewed 663 interactions (20 children × 17 referents × 2 interactions per referent), omitting 17 due to missing or incomplete data. These omissions also included instances in which the

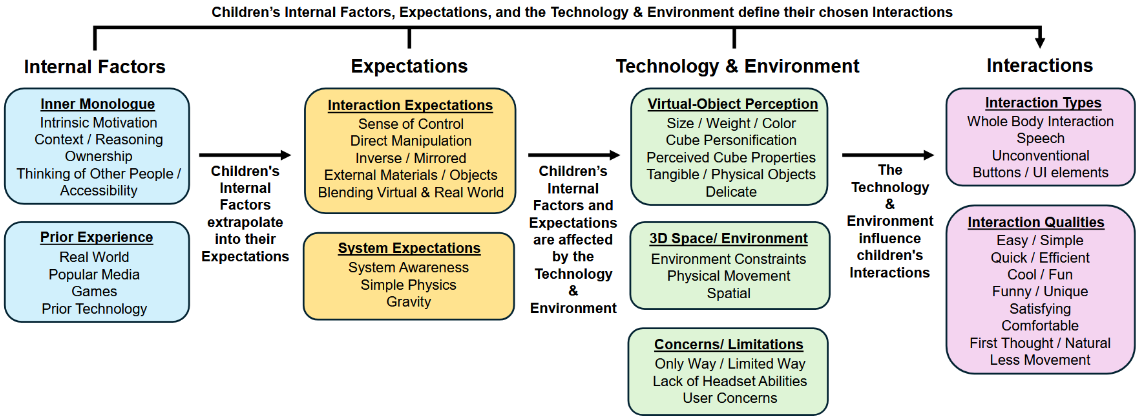


Figure 2: This figure shows our conceptual model on how children think about interacting with a virtual object in AR headsets. The model groups nine themes into four main categories (Internal Factors, Expectations, Technology & Environment, and Interactions).

child was unable to come up with an interaction. To ensure consistency in coding, the first two authors independently coded data from six children and calculated inter-rater reliability (IRR) scores across all coding categories; Cohen's kappa averaged 0.8764 (min: 0.670, max: 0.970, SD: 0.0832), corresponding to "almost perfect" agreement [75]. Following the IRR calculation, the first two authors resolved disagreements through a line-by-line consensus review. The remaining 14 children were then evenly divided between the two authors for independent coding.

To measure consensus among children's interaction proposals, we calculated the Agreement Rate ($\mathcal{AR}$), following prior methods [73, 81]. $\mathcal{AR}$ quantifies how often different participants proposed the same interaction for a given referent in an elicitation study (higher $\mathcal{AR}$ points towards increased consensus). $\mathcal{AR}$ is calculated as the ratio of the number of participant pairs who proposed equivalent interactions to the total number of possible participant pairs for that referent. Equation 1 shows the formal definition, where P is the set of all proposals for referent r, and P_i are subsets of equivalent proposals within P:

$$AR_r = \frac{\sum_{P_i \subseteq P} \frac{1}{2}|P_i|(|P_i| - 1)}{\frac{1}{2}|P|(|P| - 1)} \tag{1}$$

We used the AGATe 2.0 toolkit (Agreement Analysis Toolkit) to compute $\mathcal{AR}$ scores, referencing the source code provided by the authors [73]. We estimated 95% bootstrapped confidence intervals via percentile bootstrap (5000 resamples) for robustness, following Vatavu & Wobbrock's recommendations [74]. We defined equivalence at the level of coded proposals: two proposals agreed if they matched on interaction modality (e.g., gesture, speech, body, gesture+speech) and on the action code for that modality (e.g., push, pull, grab, tap).

4 Results

Our results are divided into two main sections: (1) the children's conceptual model, outlining the qualitative findings from our affinity diagram analysis, and (2) the interaction analysis, which presents the quantitative data and provides an overview of the children's proposed interactions.

4.1 Conceptual Model

We organized our affinity diagram into 9 themes, which we grouped into 4 main categories to capture the children's conceptual model of interacting with virtual objects in AR headsets: (1) Internal Factors, (2) Expectations, (3) Technology & Environment, and (4) Interactions (Fig. 2). The children's *Internal Factors* lie at the core of their mental models, outlining their motivations, previous experience, and creating a story as to how they want to interact with the virtual object. These internal factors extrapolate into *Expectations* that the children have about their own interactions and the system itself. Furthermore, the children's inner thought process and expectations are then affected by their understanding of the *Technology & Environment*, such as possible technological limitations. Lastly, the children's *Internal Factors*, *Expectations*, and the *Technology & Environment* define their chosen *Interactions*, which consist of a variety of types (e.g., gesture, speech, whole-body) and qualities (e.g., easy, fun, etc.). Please refer to the Data Analysis Section 3.4 for more details on our method for creating our conceptual model.

4.1.1 Internal Factors. While interacting with the virtual cube, the children verbalized their motivations, providing additional context and reasoning as part of their *Inner Monologue*. The children also drew upon *Prior Experience* as inspiration in their interactions.

Inner Monologue. We found that the children had *Intrinsic Motivations*, provided additional *Context/Reasoning* behind their interactions, displayed feelings of *Ownership* over the cube, and also were *Thinking About Other People/Accessibility* while coming up with an interaction. At the core of the children's inner thought processes lies their *Intrinsic Motivations*, which were shaped by self-set goals (e.g., having fun, wanting a challenge, collaboration). P8 (10-year-old-female) had a self-set goal to protect the cube, which they expressed throughout their interactions, such as for the Move Away referent: *"I'll push it away and I'll put it there, and then if I fall forward, I can grab the whole cube and fall, and protect the cube."*

For Move Towards, P5 (10-year-old-male) wanted to use a grappler tool that could pull the cube towards them because *"me and my brother like to play Fortnite."* The children's desire to set their own goals aligns with Self-Determination Theory, which suggests that intrinsic motivations drive human behavior [15].

Building on their intrinsic motivations, the children provided *Context/Reasoning* behind their choices, effectively creating a narrative to make sense of their actions. For instance, P14 (10-year-old-male) framed the cube as an enemy for the Make Smaller referent: *"Maybe it's like an enemy, like it's pretty threatening, you can shrink them..."* For the Destroy referent, P16 (11-year-old-male) provided additional context behind why they were deleting the cube: *"If I was building something and I didn't want it to be there, then I would say delete so it could go away."* The children's tendency to build a narrative around their interactions aligns with Fisher's Narrative Paradigm Theory [20], which proposes that humans construct stories to make sense of abstract experiences. This narrative approach may help children bridge the unfamiliar AR headset environment with familiar contexts.

Furthermore, some of the children expressed a perceived sense of *Ownership* over the virtual cube. P10 (12-year-old-female) chose to physically grab and pull the cube closer to them for Move Toward, expressing that: *"When you want something you have to hold it."* This emerging sense of ownership could be related to the children's tendency to physically manipulate the cube, as it has been studied with adults that the simple act of touching an object increases the sense of ownership over it [52]. Additionally, the children also were *Thinking About Other People/Accessibility* while creating their interactions. For Move Down, P1 decided to use speech instead of their hands: *"Because people who might wanna play and their arms are tired they could just talk."*

Prior Experience. We found that the children relied on their prior experience with the *Real World*, *Popular Media*, *Games* and *Prior Technology* as inspiration to construct their interactions. The children often used their *Real World* experience to connect everyday actions with their newly proposed interactions in the AR headset. For instance, when we asked P16 (11-year-old-male) why they grabbed and pulled the cube for Move Left, they expressed: *"It reminded me of type of sliding door where you grab it and then and you pull this way to open."* P11 (9-year-old-male) chose to cut the cube in half with their hand for Make Smaller *"Because I cut Play-Doh like that to make it smaller."* Furthermore, the children's experience with *Games* and *Popular Media*, such as movies and viral videos, influenced their interaction proposals. For Move Towards, P6 (9-year-old-male) thought of a fun way to pull the cube: *"If you are a Star Wars fan, use the force"*, while P18 (12-year-old-male) thought of slicing the cube for Destroy: *"In video games, a lot of the games use swords and I feel like it's really easy to use a sword to swipe it off."* The children's interactions were also influenced by their *Prior Technology* experience with other devices, such as computers and tablets. For Zoom In, P10 (12-year-old-female) performed a two-finger expanding gesture, explaining: *"That's how I'm used to doing it on like an iPad or something."*

4.1.2 *Expectations.* Building on their internal thought process, the children developed a number of *Interaction Expectations* and *System Expectations* which factored into what kind of interaction they wanted to propose.

Interaction Expectations. We identified that the children expected a *Sense of Control*, *Inverse/Mirrored* interactions, *Direct Manipulation*, the use of *External Material/Objects*, and *Blending the Virtual with Real World.* The children expected a general *Sense of Control* over the cube, such as P9 (11-year-old-male) using two hands for Move Left because: *"It would be controlled and there wouldn't be a chance of it falling out of my hands."* Some of the children also expected the functionality of *Inverse/Mirrored* interactions, such as squishing the cube with their hands for Make Smaller and stretching it for Make Bigger. The use of *Direct Manipulation* was consistently present across the children's interactions, with an expectation that physically manipulating the cube will result in their desired outcome.

Additionally, the children utilized *External Material/Objects* to aid with their interactions (Fig. 3). The function of the external objects varied depending on the child's goal for the interaction, such as being utilized as a tool: *"I would use a magnifying glass. Because magnifying glass is used to zoom in stuff* (P11) (9-year-old-male), or being used to gamify the interaction: *"I could use a baseball bat and hit it left...So I could play with it like a baseball"* (P1) (10-year-old-male). Some other notable examples include P8 (10-year-old-female) gluing the sides of cube together for Create, P1 and P7 (11-year-old-male) placing the cube inside a hydraulic press for the Make Smaller and Destroy referents respectively, and P2 (11-year-old-male) putting the cube in "a little train cart" and pushing it for Move Right. This expectation to use external objects sometimes led into the children *Blending the Virtual with Real World*, expecting both realities to intertwine seamlessly. For instance, P7 decided to use clay and the physical table in front of them for Create: *"I can take clay and I can kinda like form into rough cube and I can just smash against the side of the table or something and make it so that all the sides are smooth."* P11 (9-year-old-male) wanted to push the cube with a real-world stick for Move Right: *"Because I think you can grab it [the stick] from outside and then you come back in and you push it."*

System Expectations. The children expected *Simple Physics* and *Gravity* to work seamlessly in the AR environment. Sometimes these expectations were implied, such as placing a heavy object on top of the cube to move it down or using a hammer to hit the cube with increased force, and sometimes these expectations were

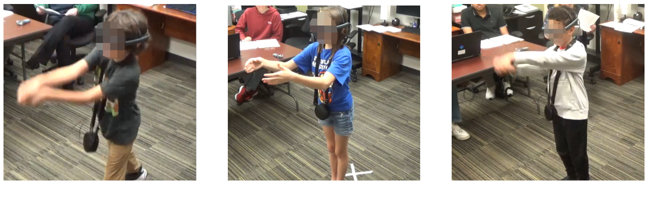

(A) (B) (C)

Figure 3: Examples of the children's interactions in which they are: (A) pretending to use a sword to destroy the cube; (B) gluing the sides of the cube together to create it; (C) pretending to hit the cube with a baseball bat to move it left.

explicit: *"If I slap it [the cube] the gravity would pull it down to the floor"* (P9, 11-year-old-male). For Make Smaller, P7 (11-year-old-male) chose an interaction for the cube they believed would make it as small as possible: *"I would have the cube on the ground, then I would have an anvil or something on a rope, suspended above, and then I would cut the rope."* In addition, some of the children expected a certain level of *System Awareness* over their actions, assuming that the system could correctly extrapolate the meaning behind their proposed interactions. For Move Right, P4 (12-year-old-female) said *"I feel like if it followed the motion of my finger then it will be like magic, it could move the cube following my finger."* For the Select Multiple referent, P10 (12-year-old-female) grabbed each cube, assuming that *"the system will know that you want that."*

4.1.3 Technology & Environment. Beyond their internal thought process and expectations, the children had their own *Virtual-Object Perceptions*, which influenced their proposed interactions. The children also took full advantage of the *3D Space/Environment*, such as proposing interactions that are only possible in 3D space. Furthermore, some of the children voiced specific *Concerns and Limitations*, which ranged from the interaction itself, such as discussing feasibility, to perceived limitations of the AR headset technology.

Virtual-Object Perceptions. Children considered the cube's *Size/Weight*, *Color*, as well as other *User Inferred Cube Properties* (e.g., the cube's material composition) to determine an appropriate interaction. For the Destroy referent, P1 (10-year-old-male) thought of putting the cube in a fire, explaining that the cube could be made out of wood and putting it in a fire *"could turn it into ash, changing the state of it."* P9 (11-year-old-male) wanted to use concrete for Create: *"Using materials or something. It [cube] looks like concrete so I would use concrete or something like that to make it"*, explaining that the cube looks like concrete because *"it's like blue-ish white so that's why I think it looks like concrete, and also it has sharp edges so I think it looks like concrete."* The children also treated the cube as a *Tangible/Physical* object, such as P6 (9-year-old-male) worrying that a foam sword would not rotate the cube, *"The cube may even break the sword."* Furthermore, some of the children went as far as considering the cube to be *Delicate*, believing that a sufficient amount of force would break it: *"So if it's [the cube] like very weak, using your hands you could smash it."* (P3, 9-year-old-male). The children's perceptions of the cube align with the Affordance Theory [22], in which users attribute affordances to objects based on their design. The cube's appearance may have prompted the children to expect their interactions to work similarly to the real world.

We also observed some instances of *Cube Personification*, in which two separate children attributed human traits to the cube. As an example, for Move Towards, P2 (11-year-old-male) decided to *"Maybe tell it [cube] we still have some pizza for lunch"* as a way to make the cube come closer to them. When asked what made them think that offering pizza would move the cube, they explained that *"Maybe it likes pizza?"* This belief that the cube has thoughts and intentions is known in child cognitive development as animism, a concept in which children attribute life-like qualities to inanimate objects [54]. Prior work examining children's expectations for intelligent touchscreen user interfaces [85] also found that children personified the system while designing by giving it human-like qualities (e.g., visual facial features).

3D Space/Environment. The children consistently employed *Physical Movement*, such as walking closer or farther away from the cube for the Zoom In/Out referents: *"Walking closer and walking away make things bigger and smaller."* (P20, 12-year-old-male). Some of the children considered the *Spatial* properties of the AR environment for their interactions. For Select One, P1 (10-year-old-male) explained that *"you can pick it up in the 3D space instead of having only click it in a 2D space."* The children also considered possible *Physical Environment Constraints* that would impede them from successfully completing their interactions, such as trying to throw the cube but not having enough physical space around them to complete the throwing motion.

Concerns and Limitations. The children expressed specific *User Concerns* about their interactions, including concerns about the cube, user ability, getting hurt, and usability. For Move Away, P8 (10-year-old-female) was concerned with falling and dropping the cube because *"if it's glass, I don't want to shatter it, and get scratched and start bleeding..."* For Zoom In, P7 decided to walk up to the cube, which raised a concern in his mind: *"...if you are on the wrong kind of ground you may trip."* There were also instances in which the children described their interaction as an *Only Way/Limited Option* to complete the referent, expressing that they could not think of another way of interacting with the cube: *"Because if you do it any other way it doesn't really look right to me"* (P12, 10-year-old-female). The children also were concerned with a *Lack of Headset Abilities*, which led them to choose interactions they believed would be easier for the system to recognize. For instance, P3 (9-year-old-male) had a strong belief that the headset would not be able to recognize any speech commands, resorting to primarily using gestures: *"because talking, like saying something, wouldn't do anything."* When we asked P3 why they believed speech would not work, they replied: *"well it's [speech] not physically doing anything, and using your body it is physically doing something."* P3 believed speech could not physically move the cube, similar to real life.

4.1.4 Interactions. The children's internal factors and expectations, combined with their understanding of the technology & environment, culminated in their proposed interactions. Their imaginative approach resulted in a diverse set of *Interaction Types* and revealed the *Interaction Qualities* they considered most important.

Interaction Types. Although gesture-based interactions were prominent (present in 76.5% of the proposals, further described in the Interaction Analysis section below), the children's interaction

(A) (B) (C)

Figure 4: Examples of children's interactions, in which they are: kicking the cube and moving it up (A) from their point of view and (B) an external camera perspective; (C) drawing the cube's 3D outline in mid-air to create it.

proposals did reflect a range of interaction types, including *Whole-Body Interaction, Speech, Unconventional Interactions,* and the use of *Buttons/UI Elements.* The children often engaged their whole body to interact with the cube (Fig. 4). For instance, P11 (9-year-old-male) described: *"I would step on it and it would go down... Sometimes I push balloons on the floor with my foot and they go down."* P7 (11-year-old-male) ran away from the cube to make it appear smaller: *"I would turn around and run wherever I need to and then look back at it. Because it's easy to do unless you are in a tight space and it's good exercise for your leg."* The children also suggested moving their heads to direct the cube's movement, expecting the cube to follow their head movement: *"If I move my head I feel like it's always going to move where my head is [moving]..."* (P13, 9-year-old-male). Interestingly, some of the children thought of using eye-based interactions, such as P15 (9-year-old-male) deciding to *"Look at it [the cube] for a long time"* as a way to select the cube.

Some children also proposed speech commands as a way to manipulate the cube. For instance, P17 (12-year-old-male) used speech for Make Bigger, explaining that *"It will be easier to do, you could pick any size you want"*, while P19 (12-year-old-male) wanted to rotate the cube with their voice because *"It's cool to tell the cube to do that'."* P6 (9-year-old-male) drew inspiration from the game Hogwarts Legacy, invoking magical spells as speech commands: *"In Hogwarts Legacy, there's a spell that makes you go up and makes you go down... 'Descendo'."*

The children also proposed *Unconventional Interactions* that went beyond expected body, speech, or gesture-based interactions. For example, P14 (10-year-old-male) snapped their fingers to make the cube teleport: *"Maybe snap in the direction and it goes in the direction that you snap your fingers... I thought of like teleportation, snap and teleport."* The children also drew inspiration from familiar actions, such as blowing on the cube to push it downward: *"Maybe the air would push it?"* (P2, 11-year-old-male), or inflating the cube with an air pump as if it were hollow: *"I would take the needle from the air pump, stick it right into the cube, and then pump air into it"* (P7, 11-year-old-male). P3 (9-year-old-male) thought of shattering the cube with a loud scream: *"If it's made of glass, you could be very loud and once you're loud enough, usually glass could break."* Prior AR gesture elicitation studies with adults also found that some participants who did not have AR experience considered using air as a way to manipulate AR objects [16].

Some of the children expected to interact with *Buttons/UI Elements* as a way to control the cube, often drawing inspiration from digital games or familiar interfaces. For example, P12 (10-year-old female) relied on conventional game symbolism, proposing plus and minus buttons to change size: *"If it was a game then a plus sign would make it bigger and a minus sign would make it smaller."* P11 (9-year-old male) suggested physical controls, such as goggles with a built-in zoom button: *"I would wear the goggles and press a button on it to zoom in."*

Interaction Qualities. The children chose interactions that were *Easy*, *Simple*, *Quick/Efficient*, and *Comfortable* to perform (i.e., the interaction did not force their hands into an awkward position). Additionally, the children wanted their interactions to be *Cool/Fun*, *Funny*, *Unique*, and *Satisfying*: *"I could smash it to destroy...It's satisfying"* (P19, 12-year-old-male). The children also relied on their *First Thought*, preferring interactions that felt *Natural* to them: *"It just feels right to do it like that, it looks pushable"* (P10, 12-year-old-female). The children highlighted these qualities as main aspects they want in their AR headset interactions, which is reflected in their high rating scores for the *Goodness of fit* ($N = 4.31/5$, $SD = 0.83$) and *Ease of use* ($N = 4.47/5$, $SD = 0.82$) statements.

Although, we saw that some of the older children (11 to 12-year-olds) recommended *Less Physical Movement* for some of their proposed interactions, they also considered *Physical Movement* in their interactions. As we discussed in *3D Space/Environment*, the children from all ages (9-12-years-old) proposed using *Physical Movement* in their interactions.

4.2 Interaction Analysis

In this section, we present the quantitative results from our dataset of 663 proposed interactions. Our goal is to provide a brief overview of these results to give context for our conceptual model. Please refer to Section 3.4 for more details on our coding process.

4.2.1 Overview. From our analysis, we found that a majority of the children's interactions were gestures (76.5%, 507/663 interactions), such as *Direct Manipulation* (e.g., grabbing, pushing, etc.), utilizing *External Material/Objects* (e.g., hitting the cube with a baseball bat), or indirect commands (e.g., pointing where the cube should move). The children also proposed combining gestures with body movement (8.0%, 53/663 interactions), such as walking and pushing the cube at the same time to move it farther away. Other modalities, such as body-only interaction (e.g., physically approaching the cube for zooming in) (5.9%, 39/663) and speech commands (5.6%, 37/663), appeared only sporadically by comparison. Notably, speech accounted for only 1.2% of first choice interactions, suggesting that the children did not see speech as a natural way to interact with the cube. The aforementioned interaction modality distribution aligns with the children's strong expectation for *Direct Manipulation* and *Physical Movement*, as observed in the qualitative findings. An analysis of the children's first and second choices revealed that 37.3% of second-choice interactions involved a shift to a different interaction modality. For example, some children started with a hand gesture and then used speech or a body-only action for their second choice. However, during most of the interaction proposals, the children did not switch interaction modalities and stayed in the same modality (i.e., gestures).

Although gesture was the preferred modality, the children's proposed interactions were highly diverse, which resulted in low consensus. On average, the first-choice agreement rate was only $\mathcal{AR} = 0.182$ (a medium level of agreement [73]), and second-choice agreement dropped to $\mathcal{AR} = 0.065$ (Table 2). Even for seemingly straightforward referents like Move Away (first-choice $\mathcal{AR} = 0.289$), the children came up with multiple distinct ways to do it (e.g., pushing the cube, throwing the cube with force, or making a shooing motion). In addition, abstract referents like Create prompted very low agreement ($\mathcal{AR} = 0.063$ for first choice), which is unsurprising given the open-ended nature of these abstract referents. We found that first-choice interactions displayed a higher agreement across most referents, compared to second-choice, with Create being the only exception ($\mathcal{AR} = 0.063$ for first-choice compared to $\mathcal{AR} = 0.096$ for second-choice). We believe two factors contributed to this agreement score discrepancy: the agreement rate depends on the number of valid proposals (Create had fewer second-choice proposals, 17 vs. 20), and several children seemed to reflect more about their second choice, which led to a slightly higher agreement rate. The children rated Create lower on *Ease of use* ($N = 3.97/5$) than the overall average ($N = 4.47/5$), indicating that the children felt Create interactions were more difficult to perform. Create's low rating for *Ease of use* aligns with its low agreement rate and open-ended nature: the children proposed different interactions and only occasionally converged on the same approach (i.e., no majority consensus). In contrast, the *Goodness of Fit* rating for Create ($N = 4.32$) was similar to the overall average ($N = 4.31$), suggesting the children felt their ideas were appropriate even if harder to perform. Overall, low consensus among the children (Table 2) underscores the creativity and individuality in the children's mental models. While the children preferred gestures, how they used them (and what narrative context they imagined) often differed.

4.2.2 Referent Groups. For translation referents (e.g., Move Up/Down), the children primarily chose gestures (72.3%, 172/238 interactions), followed by gesture combined with body movements (11.8%, 28/238 interactions), over speech commands (4.6%, 11/238 interactions). Only 2 of 120 first-choice interactions used speech (1.7%). While Move Towards yielded the highest agreement out of all referents (first-choice $\mathcal{AR} = 0.416$), the wide margin in the confidence interval (0.226-0.721) points towards the high variability in children's interactions. For 60.0% of the first-choice Move Toward interactions (12/20), the children used a grabbing and pulling motion to move the cube closer. For the remaining translation referents, proposals were diverse (e.g., punching, pushing, smacking) and agreement remained low (Table 2).

Table 2: Agreement Rates ($\mathcal{AR}$) for first and second choices, with 95% bootstrapped confidence intervals (CIs). Cells are colored by agreement level (higher is better): red for low ($\leq .100$), yellow for medium ($.100-.300$), green for high ($.300-.500$), and light-blue for very high ($> .500$) [73]; we did not see any very high agreement rates.

Referent	1st Choice $\mathcal{AR}$ (95% CI)	2nd Choice $\mathcal{AR}$ (95% CI)
Move Up	0.163 (0.095 – 0.374)	0.076 (0.064 – 0.234)
Move Down	0.132 (0.084 – 0.326)	0.064 (0.064 – 0.205)
Move Left	0.126 (0.084 – 0.316)	0.068 (0.058 – 0.216)
Move Right	0.121 (0.079 – 0.326)	0.032 (0.042 – 0.147)
Move Towards	0.416 (0.226 – 0.721)	0.063 (0.053 – 0.216)
Move Away	0.289 (0.147 – 0.563)	0.084 (0.068 – 0.232)
Destroy	0.163 (0.105 – 0.353)	0.026 (0.042 – 0.126)
Make Bigger	0.232 (0.184 – 0.416)	0.039 (0.046 – 0.170)
Make Smaller	0.189 (0.126 – 0.400)	0.070 (0.058 – 0.240)
Rotate (X)	0.111 (0.079 – 0.268)	0.026 (0.042 – 0.137)
Rotate (Y)	0.100 (0.074 – 0.274)	0.047 (0.053 – 0.163)
Rotate (Z)	0.132 (0.089 – 0.311)	0.039 (0.046 – 0.176)
Zoom In	0.163 (0.111 – 0.342)	0.124 (0.085 – 0.320)
Zoom Out	0.284 (0.189 – 0.516)	0.047 (0.047 – 0.179)
Create	0.063 (0.058 – 0.205)	0.096 (0.088 – 0.257)
Select One	0.268 (0.168 – 0.511)	0.099 (0.082 – 0.257)
Select Multiple	0.146 (0.094 – 0.351)	0.110 (0.088 – 0.279)
Overall Avg.	**0.182**	**0.065**

Among manipulation referents (e.g., Make Bigger, Rotate, Zoom), 77.3% of the interactions (211/273) were gestures. The children occasionally incorporated body movements; 8.1% (22/273) of these interactions were body-only, and another 7.0% (19/273) combined gestures with body movement. Notably, none of the children used speech interaction as their first choice, but later some of the children (8.3%, 11/133 second choices) opted for speech for their second choice as a backup strategy.

The abstract referents (e.g., Create, Destroy) saw the highest use of gestures, used during 81.6% of abstract interactions (124/152), with speech next at 9.9% (15/152). Despite a strong preference for gestures, the children's gesture proposals remained imaginative and diverse, leading to a low overall agreement rate (first-choice average $\mathcal{AR} = 0.16$). For Create, the children's ideas ranged widely, but a few patterns stood out: shaping the cube with their hands (tracing or contouring its edges), drawing a cube in mid-air with finger(s) (Fig. 4C), or assembling it from materials (e.g., lava, clay, concrete). For Destroy, the children leaned towards forceful interactions such as punching, clapping, or throwing the cube away; no interaction dominated and the agreement was low. The children proposed various interactions for Select One/Multiple, such as tapping, hovering, and even grabbing the cube(s), and agreement remained modest (first-choice $\mathcal{AR} = 0.268$ for Select One; $\mathcal{AR} = 0.146$ for Select Multiple). Notably, the percentage of speech interactions for Select One/Multiple was higher (12.0%, 9/75) than the overall average across all referents (5.6%, 37/663); although 8/9 of the Select One/Multiple speech interactions were a second choice.

4.2.3 Use of External Objects. In examining the children's proposed gesture interactions, the children often incorporated external objects as tools or props. A total of 75% of the children (15/20) used some real or imagined external object beyond their own body, in at least one interaction proposal. These external objects ranged from everyday items, like ropes and hammers, to playful or fantastical props, such as balloons or magic wands. Breaking down by referent groups, external-object use was lowest for Translation (8.8%, 21/238 translation interactions), higher for Manipulation (17.9%, 49/273 manipulation interactions), and highest for Abstract (20.4%, 31/152 abstract interactions). Compared to abstract or manipulation referents, translation referents were often carried out with simple hand motions and gestures, leading to lower overall usage of external objects. Across all referents, Create and Destroy had the most use of external objects; 29.7% of the Create proposals (11/37) and 27.5% of the Destroy proposals (11/40) used external objects. The children drew on familiar *Real World* experiences, such as crafting, building, and cutting, while blending practical utility with imagination in their interaction proposals. Create invited materials and tools for making, and Destroy prompted tools and actions tied to breaking or discarding. For Create, P6 (9-year-old male) thought of using a wand and noted, *"If it was like a wizard game and I had a wand, I would just have to (moves wand around) cube done..."* For Destroy, P6 (9-year-old male) and P18 (12-year-old male) proposed cutting the cube into pieces with a sword, P5 (10-year-old male) thought of using a trash can to get rid of the cube, and P8 (10-year-old female) proposed using soap and water, explaining they would, *"...tub it [the cube] and slowly take apart the sides..."* In manipulation referents, we saw a variety of tools and objects such as magnifying glasses (P1, 10-year-old male) for zooming, blow devices (P5, 10-year-old male) or an anvil (P7, 11-year-old male) for resizing, and weapon-like objects such as ninja stars (P6, 9-year-old male) to rotate the cube. For instance, during Make Bigger, P11 (9-year-old male) thought of using Play-Doh and Legos and mentioned, *"I would put layers around it, like Play-Doh around the walls. Because if I want to make stuff big with Legos, I put more stuff around it to make it bigger..."*

5 Discussion

We focus our discussion on 4 categories: (1) comparing our findings to prior work, (2) exploring children's use of context and narratives in their interactions, (3) analyzing our use of qualitative coding to quantify children's proposed interactions, and (4) providing a set of design recommendations for AR headset experiences for children.

5.1 Our Findings Compared to Prior Work

We found that children preferred using *Direct Manipulation* gestures, rarely considering external controllers, speech commands, or symbolic gestures (e.g., drawing an X to Destroy the cube). Our findings empirically validate Woodward et al.'s [82] conceptual model, which hypothesized that children would prioritize natural interaction (i.e., direct manipulation) in AR. By observing children's physical behavior, we confirm that this preference persists when children directly interact with an AR headset. Munsigner et al. [43] explored three AR headset interaction methods (controller, gesture, speech) through a Fitts' Law task study, which found that children rated controllers as having the highest usability and lowest fatigue compared to using gestures or speech. While children in Munsigner et al's. study found controllers to be more usable and less fatiguing, the children in our study rarely mentioned the use of controllers, pointing towards a mismatch in expectations. Furthermore, the lack of symbolic gestures proposed in our study coincides with what was observed in Rust et al.'s [61] touchscreen display study, in which children performed fewer symbolic gestures compared to adults. Connell et al. [14] conducted a gesture elicitation study analyzing children's interactions with a whole-body interface in which children proposed interactions for interacting with a virtual cube and a menu displayed on a flat-screen display. The authors also found that some children employed more body-centered gestures, similar to what we found in *Physical Movement* and *Whole-Body Interaction*. However, children in our study went beyond by also utilizing their lower body (e.g., kicking) and leveraging the 3D space offered by the AR environment (e.g., running).

In our study, only a small percentage of the children's interactions were speech (5.6%, $N = 37/663$). Lauer et al. [33] investigated the usability of AR headsets with children, in which children selected a gemstone using a tap, air-tap, or a speech command. They found that a majority of children preferred the use of speech due to its simplicity. However, Lauer et al. only provided children with three interaction options for one task (i.e., selecting) while we examined a range of tasks and elicited children's natural expectations. In just looking at selecting, we also found that children used speech more frequently compared to other tasks; however, as mentioned above, the children's use of speech for other tasks remained low. Delgado et al. [16] conducted an unconstrained AR headset elicitation study, in which adults showed how they would interact with a virtual cube. The authors found that adults sometimes preferred speech over gestures, mentioning that speech possessed multitasking capabilities, increased interaction precision, and overall simplicity. In addition, Woodward et al's. [82] online PD study reported children thought of using speech as part of their AR headset interactions; however, in their study children did not interact with an AR headset. Conversely, we found that when children directly interacted with virtual objects in an AR headset, they rarely proposed using speech in their interactions.

Children's overwhelming preference for gestures over other interaction methods, particularly their frequent use of *Direct Manipulation*, could be attributed to children constructing their mental model around interacting with the cube as if it was a real-world object. This preference for *Direct Manipulation* in interactions was also found by previous AR headset elicitation studies conducted with adults [53, 55, 78]; however, these studies constrained adults to using a single interaction type at a time, which limited the emergence of people's natural expectations. Delgado et al.'s [16] study opted for a more open-ended approach that allowed adults to better express their natural preferences. The authors found that adults would sometimes mention the use of *External Objects* as a visual aid for their interactions. Our study found that children also used *External Objects*, which even extended into children using *External Materials* (e.g., using clay to create the cube), a behavior not observed in adults [16]. We found that children used *External Objects* differently and more frequently than adults. Children often used external objects as tools, such as using a screwdriver to assemble the cube, a microscope to zoom it in, swords to push and cut, attaching a propeller to move the cube, or dropping an anvil on it so it would get smaller. Adults' use of external objects were grounded in having the object as a visual aid to explain the interaction (e.g., a separate handle for rotating the cube), not as a core component of the interaction.

In our conceptual model, children constructed their interactions based on their *Internal Factors* and *Expectations*, which were also affected by their understanding of the *Technology & Environment*. Quander et al. [57] found that children perceived AR headsets to be more intelligent compared to other devices (i.e., AR Tablet, Alexa). We found that children perceived the headset to be intelligent, leading to a variety of interaction and system expectations (e.g., *System Awareness*, *Blending Virtual and Real World*, *Simple Physics*, etc.). In comparing our conceptual model to Woodward et al.'s study [82] (mentioned above), we saw similar expectations. For example, children expected a level of system awareness from the AR headset, the virtual world to affect the real world, the use of multiple interaction types and whole-body interactions, and a concern for a lack of headset ability. While the children in Woodward et al.'s study did not physically interact with an AR headset, they found that children had expectations for system awareness and blending the virtual with real world which was consistent with what we found when children directly interacted with an AR headset. Additionally, children's use of the AR headset allowed us to see how the headset's *Technology & Environment* affected their mental model and interaction preferences.

5.2 Context and Narratives

An interesting finding from our study, uncovered in *Internal Factors*, was that children created narratives behind their interactions. These narratives were shaped by children's own *Intrinsic Motivations* and the *Context/Reasoning* they provided to support their proposed interactions. Prior AR headset elicitation studies with adults did not observe an emergence of storytelling or contextualization as core components of virtual-object interactions [16, 53, 55, 78]. In our study, children often used narratives as a prerequisite to propose interactions, giving their interaction a meaningful purpose. For Move Up, P2 (11-year-old-male) thought of kicking the cube (Fig. 4A&B), explaining *"so if I'm like walking in a tunnel or something and I try to go under it so it doesn't hit my head."*

Children's tendency to weave stories for contextualization aligns with established psychological theories. Self-Determination Theory suggests that intrinsic motivations drive human behavior [15]. We saw this in children through their self-set goals (e.g., protecting the cube, wanting a challenge), which influenced their interaction choices. The Narrative Paradigm Theory [20] proposes that humans construct stories to make sense of abstract experiences. We found that children used storytelling as a tool to connect the unfamiliar AR environment with familiar contexts. Interestingly, previous elicitation studies with children [14, 61, 66] did not remark the use storytelling in their interactions, suggesting that this behavior could have emerged due to the unique interaction possibilities offered by AR headsets, as the virtual and real world are more blended.

5.3 Qualitative Coding

We utilized qualitative coding to quantify the children's proposed interactions. While this approach is a standard technique used by prior elicitation studies with both adults and children [16, 61, 66, 68], we saw that children's unique interaction propositions in AR headsets were sometimes too complex to fully capture with a pure interaction coding approach. As an example, for Zoom Out, P1 (10-year-old-male) chose to *"fly up in a helicopter making it harder to see"* explaining that *"since helicopters can go up really high in the sky, making it harder to see stuff. You can see lots of it."* Additionally, children would often describe the same interaction theme (e.g., stretching the cube for Make bigger) but perform slightly different gestures. For instance, some children pulled opposite corners with two fingers, others grabbed and pulled the sides of the cube, and one child placed two "suction devices" on opposite sides of the cube to stretch it. Pham et al. [53] opted to focus their AR headset elicitation analysis around using interaction themes to reduce unnecessary gesture variance and to better capture how adults thought about interacting with the objects. As an example, the authors created a squish theme to group gestures that had the same intent to squish the object (e.g., clap with hand, with fist, push into ground). We recommend that future gesture elicitation studies with children and AR headsets focus on a thematic approach, as it may better capture children's diverse interactions with virtual objects.

5.4 AR Headset Design Recommendations

Based on our conceptual model and quantitative analysis, we have crafted a set of design recommendations that can provide direction for designing child-centered AR headset interactions. In each subsection, we connect the recommendation to our findings by stating the conceptual model themes that led to the recommendation, and highlight which recommendations can be currently implemented and which require additional future research. Current and future AR headset application designers can utilize our recommendations as a general guide to create more engaging and intuitive applications that could better align with children's natural expectations and preferences. Although our work provides a necessary and important foundation, more research will be needed to further understand how to create child-centered AR headset experiences (e.g., examining different contexts and tasks).

5.4.1 Support Whole-Body Interactions [Current & Future]. Our findings from *Interaction Types* and *3D Space/Environment* themes revealed that children frequently utilized their lower body (e.g., kicking to move the cube up) and locomotion (e.g., running away to zoom out) to accomplish tasks. Current AR headsets prioritize fine-motor hand gestures, such as the "air tap" or the one-handed wrist start gesture found in the Microsoft HoloLens 2 [42]. Prior work investigated the usability of Microsoft HoloLens 2 for elementary school children and found that 60% of the children disliked the air-tap gesture because it was difficult to perform.

Woodward et al. [82] suggested that future headsets should support mobility based on children's design ideas from online study sessions. However, it is important to note that in Woodward et al.'s study [82], the children did not interact with an AR headset. Our findings provide an empirical verification of Woodward et al.'s conceptual recommendation. We observed that when wearing the AR headset, children went beyond the traditional inputs supported by current hardware and wanted to utilize their entire body. Therefore, we recommend that designers focus on expanding gesture recognition libraries to better track input from the whole body, which aligns with children's interaction expectations.

5.4.2 Using External Objects to Facilitate Interactions [Current]. A key component of our *Expectations* theme was the use of *External Objects* in their interactions, a behavior that was largely absent from previous AR headset gesture elicitation studies conducted with adults [16, 53, 55, 79], as well as other studies with children and AR headsets [33, 57, 82]. One of the main reasons children opted for external objects was to use them as tools or aids that facilitated an interaction with the cube (e.g., using binoculars to zoom into the cube instead of walking towards it). Current AR headset applications often require users to learn abstract gestures to perform tasks (e.g., a specific pinch to zoom in).

Future designers of AR headset applications for children could incorporate external objects as part of a "Virtual Toolkit" in which specific gestures would summon application-specific 3D tools (e.g., a punching motion summons a virtual hammer for destruction). Furthermore, utilizing the headset's depth sensors to recognize generic physical props (e.g., a held stick acting as a wand) would bridge the child's physical environment with their virtual intent, satisfying their expectation for direct object manipulation.

5.4.3 Bridge the Virtual and Real-World [Current & Future]. Our quantitative analysis revealed that the children tended to use *Direct Manipulation* to interact with the cube, expecting that the cube will react to physical forces like a real-world object. As discussed in the

Blending the Virtual with Real World theme, some of the children expected real-world objects to interact with the virtual environment (e.g., going outside to grab sticks to push the cube), and for virtual objects to realistically interact with the real-world, believing that virtual and physical objects would be able to seamlessly interact in both environments.

Woodward et al [82] previously identified that children conceptually imagined virtual elements affecting the real world. However, the proposals from children in Woodward et al.'s study tended to be more fantastical (e.g., a virtual lightsaber cutting real cars in half), which could be due to the fact that their child participants did not actually interact with an AR headset. Our *Technology & Environment* theme revealed that while the children expected the virtual object to interact with the real world, their mental models were typically built around the idea that the cube would seamlessly interact with the real environment based on the cube's material properties. For instance, for Destroy, P9 (11-year-old-male) decided to throw the cube at the wall instead of the ground *"Because the ground is carpet, the wall is concrete, so I think it would break more easily on the wall."* Current AR experiences often treat virtual objects as floating overlays that ignore the environment [27], such as a virtual car phasing through physical walls. Therefore, we recommend that designers incorporate a context-aware physics engine that can leverage spatial mapping and computer vision to simulate realistic material interactions. If a child drops a virtual object, it should fall on top of the floor instead of phasing through it; if the object is thrown, then it should shatter or bounce off the physical surface it strikes.

5.4.4 Align Virtual-Object Appearance with Interaction Effort [Current]. In *Technology & Environment*, we found that the size, shape, and appearance of the virtual object influenced how children perceived the object and interacted with it. While prior work with AR headsets [53] found that the geometry of a virtual object (i.e., size and shape) influences adults' gesture choices, our findings uncovered that the appearance (e.g., visual textures) of the virtual object, in addition to size and shape, led children to attribute physical properties (e.g., weight, friction, and fragility) to the virtual object. For instance, P8 was concerned with dropping the cube, reasoning that it could be made out of "glass" and might shatter, while P6 worried the cube was hard enough to break a real-world foam sword. Therefore, designers should focus on connecting the virtual object's appearance with the intended action. For example, if a two-handed gesture is required, the virtual object should appear heavy (e.g., have a metallic texture, larger size) so children can more naturally interact with the object in an AR headset application.

5.4.5 Incorporate Context & Storytelling [Current & Future]. As previously discussed in the Discussion section on Context and Narratives, we found that children used storytelling to map the unfamiliar AR referents to actions and contexts familiar to them. While contextualization and storytelling can be great ways for children to construct a mental model for virtual-object interactions in AR headsets, it can also lead to interaction mismatches between children's own internal narrative and the system's functionality. Therefore, we recommend designers incorporate a narrative/storytelling component when introducing system functions and gestures to children. Teaching children how to interact with the application through a story relevant to the application's context could better inform children of the system's functionality and capabilities.

6 Limitations and Future Work

There are limitations to the scope of our work. We conducted our elicitation study with 20 children, which might seem low, but is consistent with prior work [14, 43, 61, 66]. Also, in our study, four of the children were female. While we did not break up our analysis by gender, we saw females present throughout our qualitative themes. A total of 17 children in our study had some form of prior VR headset experience, which may have biased their interactions with the virtual cube in the AR headset. Another limitation of our study is that all the children were recruited from the same city. Therefore, their economic, cultural, and societal background might be similar to each other, as we did not control for this. Future work should recruit children from geographically different areas. Additionally, we used a 3D virtual cube and 17 basic referents to provide a foundation of children's expectations of virtual object interaction. While we are able to establish a baseline, the cube and referents may also have influenced the children's interaction proposals. Additionally, we did not analyze how a different context could influence children's preferences. Future research should look at the impact of context on interaction preferences, as well as interactions with different virtual objects (e.g., vehicles, characters, game elements) and evaluate more complex referents in different interaction contexts (e.g., collaboration, productivity, gaming). Furthermore, we did not examine display factors such as field of view constraints, occlusion, or depth perception. Future work could analyze whether these display factors affect the children's expectations. Finally, beyond self-reported ratings, we did not evaluate the proposed interactions for usability, effectiveness, or cognitive workload, which future research should investigate.

7 Conclusion

Despite children increasingly using AR headsets in different contexts, little is known about how children expect to interact with virtual objects in the headsets. To design a more engaging and intuitive AR experience for children, it is important to understand how they conceptualize AR headset object interactions. In this work, we conducted an elicitation study with 20 children (ages 9-12) in which they saw a virtual 3D cube complete 17 different referents (e.g., Move Up, Destroy) and proposed interactions for the referents. To our knowledge, we are the first to explore children's expectations of virtual-object interactions through the use of gesture elicitation with an AR headset. We found that the children mostly used their hands and bodies to directly interact with the cube. Our conceptual model shows that children's internal factors (e.g., intrinsic motivations, prior experiences) shape their interaction expectations (e.g., direct manipulation, usage of external objects) and system expectations (e.g., basic physics, system awareness). Children's internal factors and expectations are then affected by their technology and environment perceptions (e.g., attributes of the virtual object, environment constraints). Together, these elements shaped the children's creative and diverse set of interaction proposals. We also present recommendations for designing intuitive AR headset applications for children that align with their conceptual models, such as using external objects to facilitate interactions, connecting interactions with real-world parallels, and incorporating storytelling to create more engaging and intuitive AR headset applications.

References

[1] Lucidchart . 2025. Lucidchart | Diagramming Powered By Intelligence.

[2] Thomas Alsop. 2024. Topic: AR Glasses and Headsets. https://www.statista.com/topics/10134/ar-glasses/.

[3] Troels L. Andersen, Sune Kristensen, Bjørn W. Nielsen, and Kaj Grønbæk. 2004. Designing an Augmented Reality Board Game with Children: The Battleboard 3D Experience. In *Proceedings of the 2004 Conference on Interaction Design and Children: Building a Community (IDC '04)*. Association for Computing Machinery, New York, NY, USA, 137–138. https://doi.org/10.1145/1017833.1017858

[4] Lisa Anthony, Quincy Brown, Jaye Nias, Berthel Tate, and Shreya Mohan. 2012. Interaction and Recognition Challenges in Interpreting Children's Touch and Gesture Input on Mobile Devices. In *Proceedings of the 2012 ACM International Conference on Interactive Tabletops and Surfaces (ITS '12)*. Association for Computing Machinery, New York, NY, USA, 225–234. https://doi.org/10.1145/2396636.2396671

[5] Christopher R. Austin, Barrett Ens, Kadek Ananta Satriadi, and Bernhard Jenny. 2020. Elicitation Study Investigating Hand and Foot Gesture Interaction for Immersive Maps in Augmented Reality. *Cartography and Geographic Information Science* 47, 3 (May 2020), 214–228. https://doi.org/10.1080/15230406.2019.1696232

[6] Valentin Bauer, Tifanie Bouchara, Olivier Duris, Charlotte Labossière, Marie-Noëlle Clément, and Patrick Bourdot. 2022. Evaluating the Acceptability and Usability of a Head-Mounted Augmented Reality Approach for Autistic Children with High Support Needs. In *International Confernece Virtual Reality and Mixed Reality (EuroXR)*. Springer International Publishing, Cham, 53–72.

[7] Hugh Beyer and Karen Holtzblatt. 1999. Contextual Design. *interactions* 6, 1 (Jan. 1999), 32–42. https://doi.org/10.1145/291224.291229

[8] Anol Bhattacherjee. 2001. Understanding Information Systems Continuance: An Expectation-Confirmation Model. *MIS Quarterly* 25, 3 (2001), 351–370. https://doi.org/10.2307/3250921 jstor:3250921

[9] Robin Brewer, Lisa Anthony, Quincy Brown, Germaine Irwin, Jaye Nias, and Berthel Tate. 2013. Using Gamification to Motivate Children to Complete Empirical Studies in Lab Environments. In *Proceedings of the 12th International Conference on Interaction Design and Children*. ACM, New York New York USA, 388–391. https://doi.org/10.1145/2485760.2485816

[10] Hacer Bülbül and Fatih Özdinç. 2022. How Real Is Augmented Reality in Pre-School? An Examination of Young Children's AR Experiences. *Journal of Theoretical Educational Science* 15, 4 (Oct. 2022), 884–906.

[11] Thomas J. Caruso, Martine Madill, Douglas Sidell, Kara Meister, Ellen Wang, Maria Menendez, Madison N. Kist, and Samuel Rodriguez. 2021. Using Augmented Reality to Reduce Fear and Promote Cooperation During Pediatric Otolaryngologic Procedures. *The Laryngoscope* 131, 4 (2021), E1342–E1344. https://doi.org/10.1002/lary.29098

[12] Nabil Al Nahin Ch, Diana Tosca, Tyanna Crump, Alberta Ansah, Andrew Kun, and Orit Shaer. 2022. Gesture and Voice Commands to Interact With AR Windshield Display in Automated Vehicle: A Remote Elicitation Study. In *Proceedings of the 14th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI '22)*. Association for Computing Machinery, New York, NY, USA, 171–182. https://doi.org/10.1145/3543174.3545257

[13] Yen-Ju Chen, Su-Fen Cheng, Pi-Chang Lee, Chi-Hsiu Lai, I-Ching Hou, and Chi-Wen Chen. 2020. Distraction Using Virtual Reality for Children during Intravenous Injections in an Emergency Department: A Randomised Trial. *Journal of Clinical Nursing* 29, 3-4 (2020), 503–510. https://doi.org/10.1111/jocn.15088

[14] Sabrina Connell, Pei-Yi Kuo, Liu Liu, and Anne Marie Piper. 2013. A Wizard-of-Oz Elicitation Study Examining Child-Defined Gestures with a Whole-Body Interface. In *Proceedings of the 12th International Conference on Interaction Design and Children (IDC '13)*. Association for Computing Machinery, New York, NY, USA, 277–280. https://doi.org/10.1145/2485760.2485823

[15] Edward L. Deci and Richard M. Ryan. 2000. The "What" and "Why" of Goal Pursuits: Human Needs and the Self-Determination of Behavior. *Psychological Inquiry* 11, 4 (Oct. 2000), 227–268. https://doi.org/10.1207/S15327965PLI1104_01

[16] Alejandro Delgado, Hetvi Shah, Md Mehedi Hasan Jibon, Anzhelika Kurnikova, Fareeza Rahman, and Julia Woodward. 2025. "Because We Are Human, We Use Hands": Understanding People's Expectations and Mental Models of Virtual Object Interaction in AR Headsets. In *Proceedings of the Extended Abstracts of the CHI Conference on Human Factors in Computing Systems (CHI EA '25)*. Association for Computing Machinery, New York, NY, USA, 1–8. https://doi.org/10.1145/3706599.3719691

[17] Haiwei Dong, Ali Danesh, Nadia Figueroa, and Abdulmotaleb El Saddik. 2015. An Elicitation Study on Gesture Preferences and Memorability Toward a Practical Hand-Gesture Vocabulary for Smart Televisions. *IEEE Access* 3 (2015), 543–555. https://doi.org/10.1109/ACCESS.2015.2432679

[18] Ze Dong, Thammathip Piumsomboon, Jingjing Zhang, Adrian Clark, Huidong Bai, and Rob Lindeman. 2020. A Comparison of Surface and Motion User-Defined Gestures for Mobile Augmented Reality. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems (CHI EA '20)*. Association for Computing Machinery, New York, NY, USA, 1–8. https://doi.org/10.1145/3334480.3382883

[19] Shariff A. M. Faleel, Michael Gammon, Yumiko Sakamoto, Carlo Menon, and Pourang Irani. 2020. User Gesture Elicitation of Common Smartphone Tasks for Hand Proximate User Interfaces. In *Proceedings of the 11th Augmented Human International Conference (AH '20)*. Association for Computing Machinery, New York, NY, USA, 1–8. https://doi.org/10.1145/3396339.3396363

[20] Walter R. Fisher. 1984. Narration as a Human Communication Paradigm: The Case of Public Moral Argument. *Communication Monographs* 51, 1 (March 1984), 1–22. https://doi.org/10.1080/03637758409390180

[21] GamesBeat. 2018. 10 Ethical Concerns That Will Shape the VR Industry.

[22] James J. Gibson. 2014. *The Ecological Approach to Visual Perception: Classic Edition*. Psychology Press, New York. https://doi.org/10.4324/9781315740218

[23] Anne-Laure Guinet, Guillaume Bouyer, Samir Otmane, and Eric Desailly. 2021. Validity of Hololens Augmented Reality Head Mounted Display for Measuring Gait Parameters in Healthy Adults and Children with Cerebral Palsy. *Sensors* 21, 8 (Jan. 2021), 2697. https://doi.org/10.3390/s21082697

[24] Nick Haber, Catalin Voss, and Dennis Wall. 2020. Upgraded Google Glass Helps Autistic Kids "See" Emotions. https://spectrum.ieee.org/upgraded-google-glass-helps-autistic-kids-see-emotions/particle-8

[25] Lynne Hall, Colette Hume, and Sarah Tazzyman. 2016. Five Degrees of Happiness: Effective Smiley Face Likert Scales for Evaluating with Children. In *Proceedings of the The 15th International Conference on Interaction Design and Children*. ACM, Manchester United Kingdom, 311–321. https://doi.org/10.1145/2930674.2930719

[26] Juan Pablo Hourcade, Summer Schmuecker, Delaney Norris, Meredith Onions, and Amy Gilhoi. 2025. Eliciting Preschool Children's Preferences for Augmented Reality Smart Glasses. In *Proceedings of the 24th Interaction Design and Children*. ACM, Reykjavik Iceland, 608–621. https://doi.org/10.1145/3713043.3728861

[27] Md Mehedi Hasan Jibon, Ngu Quoc Truong, Tanzila Roushan Milky, Felicia Rose Drysdale, and Julia Woodward. 2025. Synthesizing Evidence-Based AR Design Recommendations and Identifying Gaps in Practice. In *31st ACM Symposium on Virtual Reality Software and Technology (VRST '25)*. ACM, Montreal, QC, Canada. https://doi.org/10.1145/3756884.3766012

[28] Carmen M. Juan, Giacomo Toffetti, Francisco Abad, and Juan Cano. 2010. Tangible Cubes Used as the User Interface in an Augmented Reality Game for Edutainment. In *2010 10th IEEE International Conference on Advanced Learning Technologies*. IEEE, Sousse, Tunisia, 599–603. https://doi.org/10.1109/ICALT.2010.170

[29] Polyxeni Kaimara, Andreas Oikonomou, and Ioannis Deliyannis. 2022. Could Virtual Reality Applications Pose Real Risks to Children and Adolescents? A Systematic Review of Ethical Issues and Concerns. *Virtual Reality* 26, 2 (2022), 697–735. https://doi.org/10.1007/s10055-021-00563-w

[30] Sumbul Khan and Bige Tunçer. 2019. Gesture and Speech Elicitation for 3D CAD Modeling in Conceptual Design. *Automation in Construction* 106 (Oct. 2019), 102847. https://doi.org/10.1016/j.autcon.2019.102847

[31] Seongki Kim, JinHo Ryu, Youngchyul Choi, YooSeok Kang, Hongle Li, and Kibum Kim. 2020. Eye-Contact Game Using Mixed Reality for the Treatment of Children With Attention Deficit Hyperactivity Disorder. *IEEE Access* 8 (2020), 45996–46006. https://doi.org/10.1109/ACCESS.2020.2977688

[32] Mikko Korkiakoski, Paula Alavesa, and Panos Kostakos. 2024. Preference in Voice Commands and Gesture Controls With Hands-Free Augmented Reality With Novel Users. *IEEE Pervasive Computing* 23, 1 (Jan. 2024), 18–26. https://doi.org/10.1109/MPRV.2024.3364541

[33] Luisa Lauer, Kristin Altmeyer, Sarah Malone, Michael Barz, Roland Brünken, Daniel Sonntag, and Markus Peschel. 2021. Investigating the Usability of a Head-Mounted Display Augmented Reality Device in Elementary School Children. *Sensors* 21, 19 (Jan. 2021), 6623. https://doi.org/10.3390/s21196623

[34] Magic Leap. 2025. Product Specifications. https://www.magicleap.care/hc/en-us/articles/7813913215373-Product-Specifications.

[35] Yu Liu, Jessica L. Bitter, and Ulrike Spierling. 2023. Evaluating Interaction Challenges of Head-Mounted Device-Based Augmented Reality Applications for First-Time Users at Museums and Exhibitions. In *Culture and Computing*, Matthias Rauterberg (Ed.). Springer Nature Switzerland, Cham, 150–163.

[36] Silvia Lovato and Anne Marie Piper. 2015. "Siri, Is This You?": Understanding Young Children's Interactions with Voice Input Systems. In *Proceedings of the 14th International Conference on Interaction Design and Children (IDC '15)*. Association for Computing Machinery, New York, NY, USA, 335–338. https://doi.org/10.1145/2771839.2771910

[37] Silvia B. Lovato, Anne Marie Piper, and Ellen A. Wartella. 2019. Hey Google, Do Unicorns Exist? Conversational Agents as a Path to Answers to Children's Questions. In *Proceedings of the 18th ACM International Conference on Interaction Design and Children (IDC '19)*. Association for Computing Machinery, New York, NY, USA, 301–313. https://doi.org/10.1145/3311927.3323150

[38] leap magic. 2025. Magic Leap 2: Next-Gen Augmented Reality Headset for Professionals. https://www.magicleap.com/magic-leap-2.

[39] Nora McDonald, Sarita Schoenebeck, and Andrea Forte. 2019. Reliability and Inter-rater Reliability in Qualitative Research: Norms and Guidelines for CSCW

and HCI Practice. *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW (Nov. 2019), 1–23. https://doi.org/10.1145/3359174

[40] Meta. 2024. TTC Labs - Co-designing Immersive Social Age-Appropriate Experiences. https://www.ttclabs.net/report/co-designing-immersive-social-age-appropriate-experiences-insights-from-young-people-parents-and-advisors.

[41] Microsoft. 2023. HoloLens 2 Hardware. https://learn.microsoft.com/en-us/hololens/hololens2-hardware.

[42] Microsoft. 2025. Getting around HoloLens 2. https://learn.microsoft.com/en-us/hololens/hololens2-basic-usage.

[43] Brita Munsinger and John Quarles. 2019. Augmented Reality for Children in a Confirmation Task: Time, Fatigue, and Usability. In *Proceedings of the 25th ACM Symposium on Virtual Reality Software and Technology* (Parramatta, NSW, Australia) *(VRST '19)*. Association for Computing Machinery, New York, NY, USA, Article 36, 5 pages. https://doi.org/10.1145/3359996.3364274

[44] Brita Munsinger, Greg White, and John Quarles. 2019. The Usability of the Microsoft HoloLens for an Augmented Reality Game to Teach Elementary School Children. In *2019 11th International Conference on Virtual Worlds and Games for Serious Applications (VS-Games)*. IEEE, Vienna, Austria, 1–4. https://doi.org/10.1109/VS-Games.2019.8864548

[45] Vijayakumar Nanjappan, Rongkai Shi, Hai-Ning Liang, Haoru Xiao, Kim King-Tong Lau, and Khalad Hasan. 2019. Design of Interactions for Handheld Augmented Reality Devices Using Wearable Smart Textiles: Findings from a User Elicitation Study. *Applied Sciences* 9, 15 (Jan. 2019), 3177. https://doi.org/10.3390/app9153177

[46] Michelle M. Neumann, Meryl K. Keioskie, Dale Patterson, and David L. Neumann. 2022. Virtual, Augmented, and Mixed Reality: Benefits and Barriers for Early Childhood Education. *Childhood Education* 98, 4 (July 2022), 68–79. https://doi.org/10.1080/00094056.2022.2108298

[47] Chloe Ng and Nicolai Marquardt. 2022. Eliciting User-Defined Touch and Mid-air Gestures for Co-located Mobile Gaming. *Proc. ACM Hum.-Comput. Interact.* 6, ISS (Nov. 2022), 569:303–569:327. https://doi.org/10.1145/3567722

[48] Francisco R. Ortega, Katherine Tarre, Mathew Kress, Adam S. Williams, Armando B. Barreto, and Naphtali D. Rishe. 2019. Selection and Manipulation Whole-Body Gesture Elicitation Study In Virtual Reality. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, Osaka, Japan, 1723–1728. https://doi.org/10.1109/VR.2019.8798105

[49] Ahmed Oun, Nathan Hagerdorn, Caleb Scheideger, and Xiangyi Cheng. 2024. Mobile Devices or Head-Mounted Displays: A Comparative Review and Analysis of Augmented Reality in Healthcare. *IEEE Access* 12 (2024), 21825–21839. https://doi.org/10.1109/ACCESS.2024.3361833

[50] Gülçin Özalp Gerçeker, Dijle Ayar, Emine Zahide Özdemir, and Murat Bektaş. 2020. Effects of Virtual Reality on Pain, Fear and Anxiety during Blood Draw in Children Aged 5–12 Years Old: A Randomised Controlled Study. *Journal of Clinical Nursing* 29, 7-8 (2020), 1151–1161. https://doi.org/10.1111/jocn.15173

[51] David Passig and Timor Schwartz. 2014. Solving Conceptual and Perceptual Analogies with Virtual Reality among Kindergarten Children of Immigrant Families. *Teachers College Record* 116, 2 (Feb. 2014), 1–36. https://doi.org/10.1177/016146811411600205

[52] Joann Peck and Suzanne B. Shu. 2009. The Effect of Mere Touch on Perceived Ownership. *Journal of Consumer Research* 36, 3 (Oct. 2009), 434–447. https://doi.org/10.1086/598614

[53] Tran Pham, Jo Vermeulen, Anthony Tang, and Lindsay MacDonald Vermeulen. 2018. Scale Impacts Elicited Gestures for Manipulating Holograms: Implications for AR Gesture Design. In *Proceedings of the 2018 Designing Interactive Systems Conference (DIS '18)*. Association for Computing Machinery, New York, NY, USA, 227–240. https://doi.org/10.1145/3196709.3196719

[54] Jean Piaget. 1989. *The Child's Conception of the World: A 20th-Century Classic of Child Psychology.* Bloomsbury Academic, London, England.

[55] Thammathip Piumsomboon, Adrian Clark, Mark Billinghurst, and Andy Cockburn. 2013. User-Defined Gestures for Augmented Reality. In *Human-Computer Interaction – INTERACT 2013*, Paula Kotzé, Gary Marsden, Gitte Lindgaard, Janet Wesson, and Marco Winckler (Eds.). Springer, Berlin, Heidelberg, 282–299. https://doi.org/10.1007/978-3-642-40480-1_18

[56] Patricia Pons and Javier Jaen. 2020. Interactive Spaces for Children: Gesture Elicitation for Controlling Ground Mini-Robots. *Journal of Ambient Intelligence and Humanized Computing* 11, 6 (June 2020), 2467–2488. https://doi.org/10.1007/s12652-019-01290-6

[57] Kai Quander, Tanzila Roushan Milky, Natalie Aponte, Natalia Caceres Carrascal, and Julia Woodward. 2024. "Are You Smart?": Children's Understanding of "Smart" Technologies. In *Proceedings of the 23rd Annual ACM Interaction Design and Children Conference (IDC '24)*. Association for Computing Machinery, New York, NY, USA, 625–638. https://doi.org/10.1145/3628516.3655787

[58] Shwetha Rajaram, Chen Chen, Franziska Roesner, and Michael Nebeling. 2023. Eliciting Security & Privacy-Informed Sharing Techniques for Multi-User Augmented Reality. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23)*. Association for Computing Machinery, New York, NY, USA, 1–17. https://doi.org/10.1145/3544548.3581089

[59] Janet C Read and Matt Horton. 2025. Using the Smileyometer to Measure UX with Children. *Interacting with Computers* (May 2025), iwaf016. https://doi.org/10.1093/iwc/iwaf016

[60] Jaime Ruiz, Yang Li, and Edward Lank. 2011. User-Defined Motion Gestures for Mobile Interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11)*. Association for Computing Machinery, New York, NY, USA, 197–206. https://doi.org/10.1145/1978942.1978971

[61] Karen Rust, Meethu Malu, Lisa Anthony, and Leah Findlater. 2014. Understanding Childdefined Gestures and Children's Mental Models for Touchscreen Tabletop Interaction. In *Proceedings of the 2014 Conference on Interaction Design and Children (IDC '14)*. Association for Computing Machinery, New York, NY, USA, 201–204. https://doi.org/10.1145/2593968.2610452

[62] Ned T. Sahin, Neha U. Keshav, Joseph P. Salisbury, and Arshya Vahabzadeh. 2018. Safety and Lack of Negative Effects of Wearable Augmented-Reality Social Communication Aid for Children and Adults with Autism. *Journal of Clinical Medicine* 7, 8 (2018). https://doi.org/10.3390/jcm7080188

[63] T. Saliba, D. Schmartz, J.-F. Fils, and P. Van Der Linden. 2022. The Use of Virtual Reality in Children Undergoing Vascular Access Procedures: A Systematic Review and Meta-Analysis. *Journal of Clinical Monitoring and Computing* 36, 4 (Aug. 2022), 1003–1012. https://doi.org/10.1007/s10877-021-00725-w

[64] Shaikh Shawon Arefin Shimon, Sarah Morrison-Smith, Noah John, Ghazal Fahimi, and Jaime Ruiz. 2015. Exploring User-Defined Back-Of-Device Gestures for Mobile Devices. In *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '15)*. Association for Computing Machinery, New York, NY, USA, 227–232. https://doi.org/10.1145/2785830.2785890

[65] Kiley Sobel. 2019. Future of Childhood Immersive Media and Child Development. https://files.eric.ed.gov/fulltext/ED598949.pdf

[66] Nikita Soni, Schuyler Gleaves, Hannah Neff, Sarah Morrison-Smith, Shaghayegh Esmaeili, Ian Mayne, Sayli Bapat, Carrie Schuman, Kathryn A. Stofer, and Lisa Anthony. 2019. Do User-Defined Gestures for Flatscreens Generalize to Interactive Spherical Displays for Adults and Children?. In *Proceedings of the 8th ACM International Symposium on Pervasive Displays (PerDis '19)*. Association for Computing Machinery, New York, NY, USA, 1–7. https://doi.org/10.1145/3321335.3324941

[67] Erica Southgate, Shamus P. Smith, and Jill Scevak. 2017. Asking Ethical Questions in Research Using Immersive Virtual and Augmented Reality Technologies with Children and Youth. In *2017 IEEE Virtual Reality (VR)*. 12–18. https://doi.org/10.1109/VR.2017.7892226

[68] Giovanni Maria Troiano, Esben Warming Pedersen, and Kasper Hornbæk. 2014. User-Defined Gestures for Elastic, Deformable Displays. In *Proceedings of the 2014 International Working Conference on Advanced Visual Interfaces*. ACM, Como Italy, 1–8. https://doi.org/10.1145/2598153.2598184

[69] Lawrence Tychsen and Paul Foeller. 2020. Effects of Immersive Virtual Reality Headset Viewing on Young Children: Visuomotor Function, Postural Stability, and Motion Sickness. *American Journal of Ophthalmology* 209 (Jan. 2020), 151–159. https://doi.org/10.1016/j.ajo.2019.07.020

[70] D.W.F. Van Krevelen and R. Poelman. 2010. A Survey of Augmented Reality Technologies, Applications and Limitations. *International Journal of Virtual Reality* 9, 2 (Jan. 2010), 1–20. https://doi.org/10.20870/IJVR.2010.9.2.2767

[71] Marieke van Rooij, Adam Lobel, Owen Harris, Niki Smit, and Isabela Granic. 2016. DEEP: A Biofeedback Virtual Reality Game for Children At-risk for Anxiety. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '16)*. Association for Computing Machinery, New York, NY, USA, 1989–1997. https://doi.org/10.1145/2851581.2892452

[72] Sergei Vardomatski. 2021. Council Post: Augmented And Virtual Reality After Covid-19. https://www.forbes.com/councils/forbestechcouncil/2021/09/14/augmented-and-virtual-reality-after-covid-19/.

[73] Radu-Daniel Vatavu and Jacob O. Wobbrock. 2015. Formalizing Agreement Analysis for Elicitation Studies: New Measures, Significance Test, and Toolkit. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. ACM, Seoul Republic of Korea, 1325–1334. https://doi.org/10.1145/2702123.2702223

[74] Radu-Daniel Vatavu and Jacob O. Wobbrock. 2022. Clarifying Agreement Calculations and Analysis for End-User Elicitation Studies. *ACM Trans. Comput.-Hum. Interact.* 29, 1 (Jan. 2022), 5:1–5:70. https://doi.org/10.1145/3476101

[75] Anthony J. Viera and Joanne M. Garrett. 2005. Understanding Interobserver Agreement: The Kappa Statistic. *Family Medicine* 37, 5 (May 2005), 360–363.

[76] Rafael Villena-Taranilla, Sergio Tirado-Olivares, Ramón Cózar-Gutiérrez, and José Antonio González-Calero. 2022. Effects of Virtual Reality on Learning Outcomes in K-6 Education: A Meta-Analysis. *Educational Research Review* 35 (Feb. 2022), 100434. https://doi.org/10.1016/j.edurev.2022.100434

[77] Alla Vovk, Fridolin Wild, Will Guest, and Timo Kuula. 2018. Simulator Sickness in Augmented Reality Training Using the Microsoft HoloLens. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–9. https://doi.org/10.1145/3173574.3173783

[78] Adam S. Williams, Jason Garcia, and Francisco Ortega. 2020. Understanding Multimodal User Gesture and Speech Behavior for Object Manipulation in Augmented Reality Using Elicitation. *IEEE Transactions on Visualization and Computer Graphics* 26, 12 (Dec. 2020), 3479–3489. https://doi.org/10.1109/TVCG.2020.3023566

[79] Adam S. Williams and Francisco R. Ortega. 2020. Understanding Gesture and Speech Multimodal Interactions for Manipulation Tasks in Augmented Reality Using Unconstrained Elicitation. *Proceedings of the ACM on Human-Computer Interaction* 4, ISS (Nov. 2020), 1–21. https://doi.org/10.1145/3427330

[80] Jacob O. Wobbrock, Htet Htet Aung, Brandon Rothrock, and Brad A. Myers. 2005. Maximizing the Guessability of Symbolic Input. In *CHI '05 Extended Abstracts on Human Factors in Computing Systems (CHI EA '05)*. Association for Computing Machinery, New York, NY, USA, 1869–1872. https://doi.org/10.1145/1056808.1057043

[81] Jacob O. Wobbrock, Meredith Ringel Morris, and Andrew D. Wilson. 2009. User-Defined Gestures for Surface Computing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, Boston MA USA, 1083–1092. https://doi.org/10.1145/1518701.1518866

[82] Julia Woodward, Feben Alemu, Natalia E. López Adames, Lisa Anthony, Jason C. Yip, and Jaime Ruiz. 2022. "It Would Be Cool to Get Stampeded by Dinosaurs": Analyzing Children's Conceptual Model of AR Headsets Through Co-Design. In *SIGCHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 152, 13 pages. https://doi.org/10.1145/3491102.3501979

[83] Julia Woodward, Shaghayegh Esmaeili, Ayushi Jain, John Bell, Jaime Ruiz, and Lisa Anthony. 2018. Investigating Separation of Territories and Activity Roles in Children's Collaboration around Tabletops. *Proceedings of the ACM on Human-Computer Interaction* 2, CSCW (Nov. 2018), 1–21. https://doi.org/10.1145/3274454

[84] Julia Woodward, Zari McFadden, Nicole Shiver, Amir Ben-hayon, Jason C. Yip, and Lisa Anthony. 2018. Using Co-Design to Examine How Children Conceptualize Intelligent Interfaces. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, Montreal QC Canada, 1–14. https://doi.org/10.1145/3173574.3174149

[85] Julia Woodward, Zari McFadden, Nicole Shiver, Amir Ben-hayon, Jason C. Yip, and Lisa Anthony. 2018. Using Co-Design to Examine How Children Conceptualize Intelligent Interfaces. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3173574.3174149

[86] Julia Woodward and Jaime Ruiz. 2023. Designing Textual Information in AR Headsets to Aid in Adults' and Children's Task Performance. In *Proceedings of the 22nd Annual ACM Interaction Design and Children Conference* (Chicago, IL, USA) *(IDC '23)*. Association for Computing Machinery, New York, NY, USA, 27–39. https://doi.org/10.1145/3585088.3589373

[87] Julia Woodward, Alex Shaw, Annie Luc, Brittany Craig, Juthika Das, Phillip Hall, Akshay Holla, Germaine Irwin, Danielle Sikich, Quincy Brown, and Lisa Anthony. 2016. Characterizing How Interface Complexity Affects Children's Touchscreen Interactions. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) *(CHI '16)*. Association for Computing Machinery, New York, NY, USA, 1921–1933. https://doi.org/10.1145/2858036.2858200

[88] Romy Yun, Emily M He, Michelle Zuniga, Nan Guo, Ellen Y Wang, Florence Ho, Molly Pearson, Samuel T Rodriguez, and Thomas J Caruso. 2024. Augmented Reality Improves Pediatric Mask Induction: A Prospective, Matched Case-Control Study. *Journal of Patient Experience* 11 (Nov. 2024), 23743735241241146. https://doi.org/10.1177/23743735241241146

[89] Xiaoyan Zhou, Adam Sinclair Williams, and Francisco Raul Ortega. 2022. Eliciting Multimodal Gesture+Speech Interactions in a Multi-Object Augmented Reality Environment. In *Proceedings of the 28th ACM Symposium on Virtual Reality Software and Technology (VRST '22)*. Association for Computing Machinery, New York, NY, USA, 1–10. https://doi.org/10.1145/3562939.3565637