

Cite as: Dall'Olio L., Amrein O., Gianettoni, L., & Martarelli, C. S. (2024). The impact of fantasy on young children's recall: A virtual reality approach. *Virtual Reality*, 28. <https://doi.org/10.1007/s10055-023-00911-y>

The Impact of Fantasy on Young Children's Recall: A Virtual Reality Approach

Lucas Dall'Olio¹, Olivier Amrein¹, Lavinia Gianettoni², & Corinna S. Martarelli¹

¹ Faculty of Psychology, UniDistance Suisse, Switzerland

² Institute of Psychology, University of Lausanne, Switzerland

Correspondence concerning this article should be addressed to Lucas Dall'Olio (lucas.dallolio@unidistance.ch) and Corinna Martarelli (corinna.martarelli@fernuni.ch).

Funding: This study was supported by the Swiss National Science Foundation (grant number 100019_185477 awarded to C.S.M.).

Open Practices Statement

All data analyzed during the current study are available on OSF (<https://osf.io/gxvmt/>), and the study was preregistered (<https://osf.io/g2x9v>).

Acknowledgments:

We would like to thank Noëmi Capdevila, Philae Cretin, Sina Jossen, and Lilian Rau for helping with the study design and data collection. We would also like to thank all the children, as well as their parents, teachers, and school directors, who were involved in this project.

Competing Interests:

The authors declare that no competing interests exist.

Abstract

Educational materials for preschool children are often embedded with fantastical elements. Nevertheless, there is still little empirical evidence on their effectiveness, especially as concerns long-term retention. Virtual reality offers new ecological possibilities for investigating this type of learning, especially through the impact of immersion. In a between-design study, 168 children aged four to six years followed a virtual reality presentation on China presented by either a realistic young girl or an anthropomorphic animal. The level of immersion was manipulated, with half of the children following the presentation in immersive virtual reality (IVR) and the other half in desktop virtual reality (D-VR). Participants were asked to complete a new/old recognition task and a quiz task immediately after the presentation and once again one week later, with an additional transfer task being added for the second series. One week later, children performed significantly better on the new/old recognition task in the realistic condition when compared to the anthropomorphic condition. However, there were no differences observed in the quiz task and in the transfer task. It therefore seems that, under certain conditions, children remember a cultural presentation better when it is presented by a realistic avatar. The results further showed that the children performed significantly worse in the IVR conditions on all tasks. A possible explanation for this result is that IVR demands excessive cognitive resources from preschool children. Further studies should explore this unexpected result, as well as what could be done to make IVR effective for learning in preschool children.

Keywords:

Fantasy, virtual reality, recall, preschool children

Highlights

- We study the impact of fantasy and immersion on young children's recall.
- Memory consolidation was better for information presented by a realistic avatar.
- Children in desktop conditions outperformed children working with headsets.

The Impact of Fantasy on Young Children's Recall: A Virtual Reality Approach

Introduction

Whether through symbolic games, stories, television, or video games, children spend a significant amount of time in contact with fictional worlds and fantastical stories. In a systematic review of a selection of popular books, movies, and TV series for children aged three to six (i.e., top-selling books, shows with best TV ratings, movies with strong ticket sales, and movies and TV shows rented or streamed in 2016), Goldstein and Alpers (2020) found that 91.6% of these selected media contained fantastical elements. In educational settings, too, learning material is often integrated into a fantastical context. In their review of a sample of 252 educational books and videos, Chlebus et al. (2022) found that 81% contained at least one fantastical element. Anthropomorphic representation of characters is a frequent way of incorporating fantastical elements in educational materials. According to Goldstein and Alpers (2020), this is the fantastical element most often used across all types of media, accounting for 69.5% (81.4% in the case of TV shows). Despite the large use of anthropomorphic characters in recreational and educational settings, it remains somewhat uncertain whether they benefit young children's learning. Investigating whether fantastical elements, and more specifically anthropomorphism, impact preschool children's learning is the primary goal of this study.

By allowing high levels of control with high psychological relevance (Blascovich et al., 2002), virtual reality (VR) is ideally suited to investigate our research question. For several years, immersive VR devices, such as VR headsets, have been offering a new way to explore different issues related to learning and memory, with increased ecological validity and logistic ease (Smith, 2019). Moreover, with VR, it is possible to dynamically simulate not only realistic settings but also fantastical elements in a realistic manner, thus allowing for the investigation of how fantastical elements impact learning. Nevertheless, there are still mixed results about the advantages of immersive VR versus non-immersive VR learning (e.g., Babu et al., 2018; Buttussi & Chittaro, 2018; Makransky et al., 2019; Makransky et al., 2022; Olmos-Raya et al., 2018; Parong & Mayer, 2018). One of the main advantages of VR headsets compared to more traditional media, such as television or tablets, is their ability to completely immerse the user in a new world (Slater, 2003). We think that this immersive aspect could influence the perception of the realism of fantastical elements. Indeed, being immersed in a highly immersive environment could make the fantastical elements that surround children more realistic. To our knowledge, no study has addressed this issue in preschool children. Investigating the impact of immersion and the possible combined impact of immersion and realism on learning is the secondary goal of our study.

Children Learning with Fantasy

Fantastical elements are often present in the educational materials of young children. Nevertheless, the impact of such materials on children's learning and their effectiveness have not been clearly attested. Indeed, the current findings are in some measure contradictory. While researchers agree that young children can learn from fantasy (e.g., Weisberg & Hopkins, 2020; Hopkins & Lillard, 2021; Hopkins & Weisberg, 2021), there is also evidence of fantasy hampering learning, especially when it comes to the transfer of knowledge (e.g., Bonus & Mares, 2019; Richert & Smith, 2011; Walker, Gopnik, & Ganea, 2015).

One of the arguments made for the effectiveness of fantastical elements in learning is the way in which these elements may capture children's attention and thus help them stay engaged with the learning material. Hopkins and Weisberg (2020), for example, observed that children learned science-related educational content better with a story containing improbable events that violated the laws of physics than with an ordinary story. This effect was found to be even more important when the violation concerned a domain in which the children had knowledge, which in turn would indicate to them to remain particularly vigilant about what they were told (Hopkins & Weisberg, 2021). Beyond simply capturing children's attention, these elements may even induce deeper processing of the to-be-learned material (Hopkins & Lillard, 2021).

However, learning with fantastical material can be counterproductive when it comes to transferring what has been learned in an unrealistic context to the real world, especially in the long run (Bonus & Mares, 2019). Indeed, when learning with fantastical content, children must evaluate whether what they are learning is true only in the world of the story, or whether the new information is also applicable to the real world. This evaluation is called the "reader dilemma" (Hopkins & Weisberg, 2016). In a study by Richert and Smith (2011), it was more difficult for 3.5- to 4.5-year-olds to transfer the solution from a story to a problem to the real world when the story involved a fantastical character (i.e., a giant robot) than when it was a realistic character (i.e., a human being).

Different types of learning are possible through fictional content (e.g., science, history, geography). Among these, the effectiveness of cultural learning has often been investigated (Bonus & Mares, 2019; Borzekowski & Macha, 2010; Cole et al., 2003; Goldstein et al., 2022). One of the main challenges young children face in this type of learning is to correctly characterize the learning of a new culture, very different from their own, as something that can really exist rather than merely invented for the story (Mares & Sivakumar, 2014). For example, in a study by Bonus and Mares (2019), three-to-five-year-olds watched a short nine-minute video of *Sesame Street* that described aspects of Hispanic culture. They were then asked to answer comprehension questions and judge the realism of the educational content. A

week later, they were asked to prepare a party for a friend from a Hispanic country (and think about what food should be prepared, what music should be played, etc.). The authors observed that the success or failure of this transfer depended not only on memory but also on children's judgment of the realism of the video. Better memory predicted better performance on the transfer task, but only for children who remembered the content as realistic. The authors explained the underlying mechanism by arguing that if children consider educational material to be unreal, they might not integrate it into their knowledge about the real world. Another explanation is that children might invest less effort in processing material that they consider unreal (Bonus & Mares, 2019). Some studies have also shown that the number of fantastical elements in a learning material is related to children's learning performance (Geerdts, 2016; Richert & Schlesinger, 2022). It seems that many fantastical elements tend to negatively impact children's learning; thus, it has been proposed that a learning material should contain some fantasy to attract children's attention and motivate them to learn, but not too much to avoid overwhelming the children with information that they are not able to process (Weisberg et al., 2022).

Children Learning with Anthropomorphic Characters. Anthropomorphism, which involves attributing human-specific characteristics to nonhuman animals or objects (Geerdts, 2016), is a prominent approach in incorporating fantastical elements into educational materials (Goldstein and Alperson, 2020). The effect of anthropomorphism on children's learning is a subject that has yielded mixed results in research. For instance, Larsen and colleagues (2018) reported that children had more difficulty understanding the moral lessons conveyed in educational materials featuring anthropomorphic characters compared to human characters. Furthermore, studies have suggested that anthropomorphism may have a negative impact on children's learning by increasing the likelihood that they will apply anthropomorphic qualities (e.g., the entity can talk, the entity can be sad) to real-world entities (Conrad et al., 2021; Ganea et al., 2014; Li et al., 2019). On the other hand, some research did not find a significant negative effect of anthropomorphism in learning materials (Conrad et al., 2021) or has observed that it could, in some cases, enhance children's learning outcomes (Bonus and Mares, 2018; Geerdts et al., 2016).

The mixed effects of anthropomorphism, similar to the more general impact of fantasy elements, have been linked to their more or less effective integration within educational materials. Researchers have noted that a crucial factor is the connection between fantastical elements and the educational goals, as poorer learning outcomes were observed when these elements were not directly linked to the educational goals (Richert and Schlesinger, 2017). Additionally, the degree of anthropomorphism has been shown to influence the quality of learning. Excessive

anthropomorphism may overstimulate children and distract them from the learning content, while insufficient anthropomorphism may fail to capture their attention and focus (Bonus & Mares, 2018; Conrad et al., 2021; Geerdts, 2016). However, it has been highlighted that the use of anthropomorphism in books and TV shows is not independent of the explicitness of educational content (the higher the anthropomorphism, the higher the explicitness), and this confound could also account for the mixed results (Nguyentran & Weisberg, 2023).

Learning with Virtual Reality

New technologies, such as VR, are occupying an increasingly important place in our society. VR can be defined as a computer-generated environment that simulates and/or reproduces various aspects and features of the physical world (Araiza-Alba et al., 2021; Makransky & Lilleholt, 2018). The arrival of these technologies in the world of education through the digitization of educational tools and materials was inevitable. However, there is still a lack of empirical evidence on how they affect children's learning and memory.

In the cognitive theory of multimedia learning (CTML), Mayer (2009) proposed three types of cognitive processes that an adult individual must deal with during multimedia learning: essential processing, generative processing, and extraneous processing. Essential processing is the cognitive processing essential to the mental representation of the learning material. Generative processing allows one to make sense of the learning material and to understand it, and *extraneous processing* is any cognitive processing that does not support the learning objective and that, on the contrary, can actually distract the learner. Given that processing capacities are limited, one of the goals of VR should be to promote generative processing without adding extraneous processing.

It has been suggested that VR can promote generative processing by providing a realistic experience of being in a particular situation (Slater & Wilbur, 1997). VR devices are usually differentiated according to the degree of immersion they offer. Experiencing content—whether it is a game or a simple presentation—on a desktop computer is referred to as low-immersion VR or desktop VR (D-VR). In contrast, what is referred to as high-immersion VR or immersive VR (IVR) is when content is experienced with a VR headset. A sense of presence is supposed to help the user actively engage in the learning process and to improve learning outcomes. It has been shown that immersion can have a positive impact on an individual's sense of presence, as well as on the ease of learning (Makransky et al., 2022).

Researchers have observed that learning in IVR is as effective (e.g., Buttussi & Chittaro, 2018; Leder et al., 2019), and sometimes more effective (e.g., Babu et al., 2018; Olmos-Raya et al., 2018; Rupp et al., 2019) than learning

with D-VR. Among the arguments put forward in favor of learning with IVR, both the immersive and motivational aspects of IVR have been advanced, as well as the interactive experience it offers. For example, Makransky and Mayer (2022) investigated the impact of a virtual field trip to Greenland in a sample of middle school students and found that IVR outperformed D-VR in the aspects of presence, enjoyment, interest, and retention in immediate and delayed follow-up tests.

However, other studies have shown that IVR can also be less effective than D-VR when learning (Makransky et al., 2019; Parong & Mayer, 2018). Among the arguments explaining the lower efficiency of IVR, researchers have pointed to cognitive overload due to the addition of extraneous processing, which devalues the gain in generative processing (Makransky et al., 2019). In this study, extraneous processing was mainly due to the overload of information to be considered during the IVR experiment and the participants' lack of experience with IVR. Indeed, VR headsets are still relatively new for the wider population for whom participation in these experiments is also the first contact they have with IVR. Meyer et al. (2019) have shown that this extraneous processing can be reduced when the participants train and become accustomed to VR headsets. For IVR to be an effective learning medium, the learning experience must be specifically designed to take advantage of this medium (Makransky & Mayer, 2022).

Finally, most studies concerning the effectiveness of learning in IVR have focused on an adult population. To our knowledge, no study has yet investigated the use of IVR in culturally anchored learning in young children. One reason for this may be that current IVR devices have been designed with adults and young adults in mind as primary users and are not necessarily suitable for younger children. Before extending their use to a younger audience, it would be essential to observe their effects and effectiveness in the learning of such a population.

Present Study

The main goal of this study was to investigate children's ability to learn from a virtual presentation according to its realistic (vs. anthropomorphic) characteristics. We chose to investigate children between the ages of four and six because this is a key age in the development of children and their relationship with fantasy. Indeed, from the age of four, a young child can differentiate between a fantasy character and a real person (Woolley & Cox, 2007). Nevertheless, this ability still develops over the course of young children's lives, and children of about five years old will have an easier time making this distinction than children of three and four years old (Corriveau & Harris, 2015; Martarelli & Mast, 2013; Martarelli et al., 2015; Sharon & Woolley, 2004). In the present study, we manipulated the realism of the presentation according to the appearance of the avatar giving the presentation. In the realistic condition,

the presentation was made by an avatar with the appearance of a young girl. In the anthropomorphic condition, the presentation was made by an avatar with the appearance of an anthropomorphic animal (goat). The second goal of this study was to investigate the possible impact of immersion on realistic/anthropomorphic manipulation, as well as the overall impact of immersion on learning. We manipulated the immersion of the presentation via the medium: half of the children followed the presentation with an Oculus Quest 2 VR headset (IVR condition) and the other half with a tablet (D-VR condition).

The presentation that the children had to follow concerned China and its culture. We opted for culturally anchored learning because the contextual immersion that VR offers can be particularly favorable to this type of learning (Makransky & Mayer, 2022). Additionally, we chose cultural learning because it is prominently featured in educational programs for preschool children, such as Sesame Street, and has thus been extensively investigated, yielding promising results (Bonus & Mares, 2019; Borzekowski & Macha, 2010; Cole et al., 2003; Goldstein et al., 2022). To assess learning and memory, the children completed two tasks, a new/old recognition task and a quiz task, immediately after the presentation and one week later. To specifically measure the transfer of the information acquired during the virtual presentation into the real world, the children were also asked to complete a transfer task one week after the presentation. Finally, we controlled for verbal comprehension, theory of mind abilities, perception of the avatar, age, and gender. In agreement with Bonus and Mares (2019), we hypothesized that children in realistic conditions might outperform children in anthropomorphic conditions. We could only speculate about the impact of immersion on learning because of the mixed findings in the literature (e.g., Babu et al. (2018) show an advantage of IVR when compared to D-VR, whereas Makransky et al. (2019) show an advantage of D-VR when compared to IVR). Given the advantage of contextualizing the learning of IVR (Makransky & Mayer, 2022), we expected immersion to have a positive impact on the learning outcomes of the children. Moreover, when realism and immersion interact, we expected children to learn better with the realistic avatar in IVR because immersion makes the experience more realistic.

Methods

Participants

The required sample size and exclusion criteria were preregistered on Open Science Framework (OSF, <https://osf.io/g2x9v>). An a priori G*Power analysis revealed that we needed a sample of at least 128 participants to detect a medium effect (2×2 between-subjects analysis of variance; parameters: $f = 0.25$, $\alpha = 0.05$, $1-\beta = 0.80$; Faul et

al., 2007). We increased the sample size to around 40 children per group ($N = 168$, 85 female, $M = 5.4$, ranging from 4–6 years) to make sure to have enough participants after data exclusion.

Children were recruited from different preschools across Switzerland and were semi-randomly assigned to one of our four different conditions, which included a presentation by an anthropomorphic avatar in IVR ($N = 46$, 25 female, mean age = 5.39, 31 French-speaking children), a presentation by an anthropomorphic avatar in D-VR ($N = 39$, 15 female, mean age = 5.33, 31 French-speaking children), a presentation by a realistic avatar in IVR ($N = 47$, 24 female, mean age = 5.34, 33 French-speaking children), and a presentation by a realistic avatar in D-VR ($N = 36$, 17 female, mean age = 5.56, 28 French-speaking children). Since we tested different children from the same classroom on the same day, we decided for logistical and practical reasons to test each classroom in the same condition. Switzerland is a country with several national languages; 123 children were tested in French, and 45 in German. The different tools and tests were adapted accordingly, either with the official translation or adaptation by a native speaker. We obtained ethical approval for the study from the institute's ethics committee, and written parental consent was obtained for each child.

Design

In this study, we manipulated the realism of the avatar in realistic and anthropomorphic conditions, as well as the level of immersion in IVR and D-VR conditions. This gave us a 2×2 design with the conditions of an anthropomorphic presentation in IVR, an anthropomorphic presentation in D-VR, a realistic presentation in IVR, and a realistic presentation in D-VR. We measured memory performance with two tasks (old/new recognition task and quiz task) and generalization with a transfer task. Control tasks were also included to assess the children's theory of mind and language abilities. We controlled the theory of mind for its impact on the understanding of fantasy (Martarelli et al., 2015) and language abilities because of the verbal skills needed by the children during the experiment (see materials below). For exploratory reasons, we included five items to assess the perception of the avatar, as well as enjoyment. We assessed these items to control for children differentiating effectively between the two avatars as being more or less realistic (manipulation check) and to find out whether there were conditions in which they experienced more enjoyment.

Materials

Virtual Presentation and Virtual Environment

The virtual environment was designed in Unity (version 2019.4.32). Different 3D models were created using Blender software (www.blender.org) in versions 2.82 to 2.93. The textures used came from the open-source database

“TextureHaven” (www.TextureHaven.com) and were edited with Photoshop (Creative Cloud, 2019). The music and different sounds used came from a license-free database (www.freesound.org). Apart from the resolution (1833 x 1920 for each eye in IVR and 1080 x 1920 in D-VR) and the field of view ($89^{\circ} \pm 4$ in IVR and 87.1° in D-VR), the virtual environment was identical in IVR and D-VR. A video of the virtual presentation is available online on OSF (<https://osf.io/gxvmt/>).

Figure 1

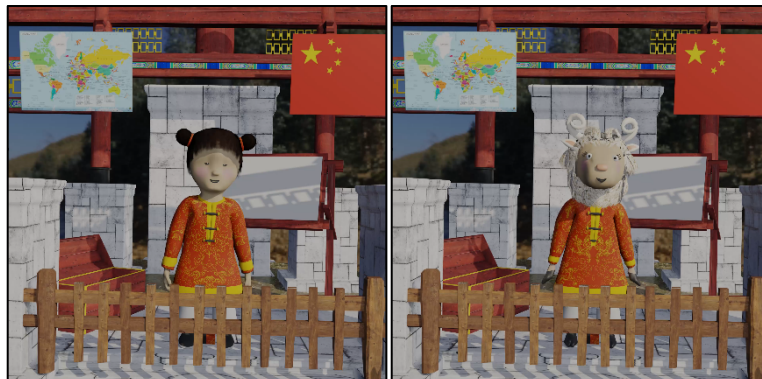
A Screenshot of the Virtual Environment



The virtual presentation consisted of a short, two-minute description of China and certain aspects of its culture by an avatar. The virtual environment consisted of a representation of a Chinese temple in a forest (see Figure 1). The virtual presentation was made either by an avatar with the appearance of a young Asian girl or by an anthropomorphic animal (goat), according to the testing condition of the participant (see Figure 2).

Figure 2

Human Avatar (Left) and Anthropomorphic Avatar (Right)



Before the presentation, each participant had time to get used to the virtual environment and the viewing medium (approximately two minutes). At the beginning of the presentation, the children were told that a character was going to appear and that they had to listen carefully to what she said and observe what she showed. During the first part of the presentation, different types of information, such as the name of the character and the location of China on the world map, were revealed. During the second part of the presentation, the avatar displayed a series of 12 visual stimuli for the participant (see Appendix A for an illustration of the stimuli). The first six stimuli were 3D models related to the food that one can typically eat in China, three others were images of typical places to visit in China, and the last three visual stimuli were images of activities to do in China. Each of these stimuli was accompanied by a commentary from the avatar that presented the 3D models and images in more detail. The different visual stimuli were either created by the authors or obtained from a copyright-free database (www.pixabay.com). The presentation was completely dubbed with the voiceover of two native speakers for our two different versions—one in German and one in French.

Previous Knowledge of China

To assess previous knowledge about China, we asked the children about the color of China's flag and the position of China on a world map. More precisely, we showed the children the Swiss flag (i.e., Do you know what this is?), described the Swiss flag, and then asked the children about the Chinese flag (color and form). With the map, the procedure was similar. We showed a map as well as the position of Switzerland on the map, and then asked the children about the position of China on the map.

Memory Tasks

To assess the children's recall of the virtual presentation, we used three different tasks: a new/old recognition task, a quiz task, and a transfer task.

New/Old Recognition Task. Designed on Psychopy (version 2020.2.10; Peirce et al., 2019), the new/old recognition task was completed on a computer by the child. This task consisted of the presentation of a series of 12 stimuli in random order, including six old stimuli and six new stimuli. The stimuli are reported in Appendix A. Old stimuli were taken from the 12 visual stimuli presented during the virtual presentation. The other 12 new stimuli were images from the internet related to things one could visit or do in another foreign country (India). For each trial, the child had to indicate whether the item was old or new by pressing a key on the computer keyboard. The task was self-paced. To facilitate the child's task, colored stickers were applied to the corresponding keys on the computer. To avoid any

confusion, the children were told that an "old" item was an image seen in the presentation, and a "new" item was an image that did not appear in the presentation. Between each trial, a fixation cross appeared for 800 ms. The child completed this task twice: once immediately after the presentation, and the second time one week later. We used two different sets (sets A and B) of 12 pictures for a total of 24 items. The order in which these sets were used (post-test vs. follow-up test) was counterbalanced.

Quiz Task. The quiz task consisted of a series of eight questions related to the presentation that the child had just completed (e.g., What was the name of the character in the story? What does her first name mean?). The questions are reported in Appendix B. Given the young age of the children, the questions were asked orally by the experimenter. Four of these questions were related to information given only orally by the avatar during the presentation (see Appendix B), while the other four were presented orally and visually during the presentation. This was done to ensure different levels of difficulty and to avoid ceiling effects. When children were unable to provide the correct answer, they were given four choices (one correct, and three wrong), and they could choose their answer among these options. Each participant completed this task twice (same questions): the first time after the presentation, and the second time one week later.

Transfer Task. For the transfer task, we presented the children with a picture of 10 different types of food on a table (see Appendix C). Among these 10 items, five were Chinese foods from the virtual presentation, and five were foods that did not come from China. We asked the children to indicate four types of food that were typical of China. The children completed this task only once (one week after the presentation).

Control Tasks

Theory of Mind Abilities. To measure the theory of mind abilities, we presented the children with four different theory of mind tasks, as used by Martarelli et al. (2015). These included two false belief tasks with representational change—the "Crayon Box" and the smarties box (Perner et al., 1987)—and two other classic theory of mind tasks—"Maxi and the Chocolate" (Wimmer & Perner, 1983) and "Mouse and the Cheese" (Clement & Perner, 1994). These tasks were enacted by the experimenter for the children using test materials (i.e., colored boxes, puppets/dolls, and wooden objects). Each task was scored as either successful or unsuccessful for a total of four points. The higher the values, the better the performance. The two false belief tasks with representational change were successfully passed if the children answered the four different questions correctly (two control and two test questions), and the two classical false belief tasks were successfully passed if the children answered the three different questions

correctly (one control and two test questions). The complete scripts for these tasks can be found on OSF (<https://osf.io/gxvmt/>).

Language Abilities. To assess language abilities, we used the corresponding subtests of the Wechsler Preschool and Primary Scale of Intelligence IV (WPPSI-IV) on active and passive language skills (Wechsler, 2012). As the French and German tests had a different number of items, we decided to use the percentage of correct answers rather than the original scoring system. For the rest of the procedure, the original instructions were respected. Language abilities were tested on the second day of testing. The two scores were averaged to create a single language abilities score. The higher the values, the better the performance.

Exploratory Items

We assessed the perception of the avatar with three items, including one about the appearance of the avatar (item: “What did the character in the video look like?”, to be answered by choosing between human and animal) and the veracity of the information she transmitted (item: “Do you believe what the character told you? Did the character tell the truth?”, to be answered by a yes/no response), as well as its possible existence in the real world (item: “Can a character like this really exist in our world? Could we meet a character like this in our world?”, to be answered by a yes/no response). Furthermore, we asked participants whether they liked the virtual presentation (item: “Did you like the virtual presentation?”). For a response, we used a pictorial measure to be answered on a five-point Likert scale (1 = not at all to 5 = a lot). Finally, following Bonus and Mares (2019), we asked the children whether they thought the presentation was just for fun or for educational purposes (choice between two options).

Procedure

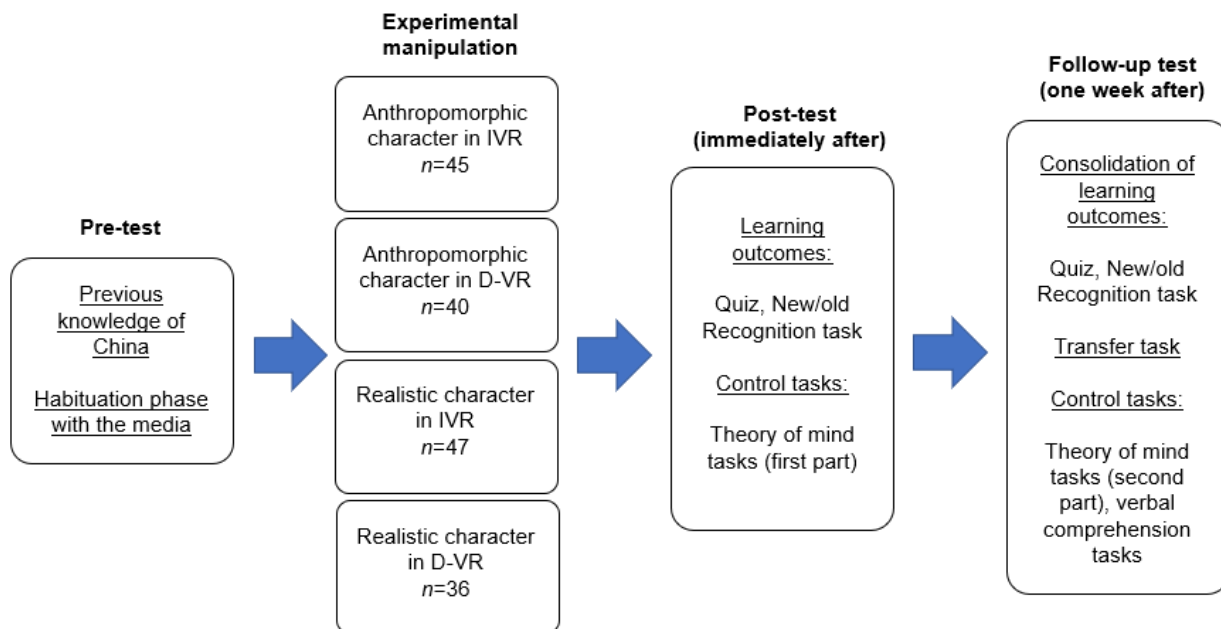
For each child, the study was separated into two 20-minute periods on two days, one week apart. The children were tested individually in isolated rooms in their respective schools. On the first day, the children had to answer a questionnaire regarding their previous knowledge about China. They were then asked to attend to a two-minute presentation about China delivered by a virtual character (young Asian girl or anthropomorphic animal) with a VR headset or a tablet, depending on their experimental condition. Before the presentation, the children had a short time (approximately two minutes) to get used to the virtual environment and the medium used (VR headset/tablet). The presentation in IVR was rendered on an Oculus Quest 2. The tablet used during the D-VR presentation was a Samsung Galaxy Tab A7. After the presentation, the children had to answer the quiz task as well as the new/old recognition task to assess what they had retained from the presentation, as well as five exploratory items to assess how they perceived

the virtual character delivering the presentation and whether they enjoyed the presentation. Finally, they had to complete the first two tasks to assess their theory of mind abilities. These tasks were carried out in a fixed order, as listed above.

One week later, the children were again asked to perform the two memory tasks, which included the quiz task and the new/old recognition task, to assess memory consolidation. After the memory tasks, the children had to perform the transfer task and finally finish the testing with the last four tasks: two assessing their theory of mind abilities, and two assessing their language abilities (passive and then active language). The study also included a questionnaire sent out to the parents that was not considered in this paper. See Figure 3 for a graphical representation of the experimental procedure.

Figure 3

Graphical Representation of the Procedure



Analytical Approach

Following our preregistration on OSF (<https://osf.io/g2x9v>), we computed different two-way analyses of variance to compare the different memory measures among the four conditions, which included the presentation by an anthropomorphic avatar in IVR, by an anthropomorphic avatar in D-VR, by a realistic avatar in IVR, and by a realistic avatar in D-VR. We report the analyses of variance in the main text and the analyses of covariance (with the covariates of age, gender, theory of mind abilities, and language skills) in the Appendix D, E and F. In addition, we computed

logistic regressions to analyze the answers to the four questions about avatar perception and the purpose of the virtual presentation. For the liking score, we computed a further analysis of variance. The logistic regressions, as well as this last analysis of variance, were not preregistered. The analyses were computed using jamovi version 1.6.23 (jamovi project, 2021). The dataset is available on OSF (<https://osf.io/gxvmt/>).

Results

Due to dropouts and technical problems, we did not have the same number of participants on the first and second days one week later; therefore, we computed analyses for each day separately. Additional analyses that consider both days in the same model, using a mixed design and focusing on a smaller sample due to the dropouts produced the same results in terms of significance and are reported in Appendix G and H. Moreover, we excluded children who had previous knowledge about China (children with a score > 2 on the basic knowledge task, $n = 4$) and outliers that were three standard deviations above/below the mean. We provide more information about data exclusion for each analysis in their respective parts.

New/Old Recognition Task

To correctly assess sensitivity, we used signal detection theory (Macmillan & Creelman, 1991). We calculated d' as in Martarelli and Mast (2013):

$$d' = z^{-1}(\text{Hit}) - z^{-1}(\text{False Alarms})$$

Where z^{-1} is the inverse of the standard normal function, Hit is the proportion of stimuli correctly categorized as old, and False Alarms are the proportion of stimuli erroneously categorized as old. As we obtained extreme values for the hit and false alarm rates (1 and 0, respectively), we adjusted these values according to the method proposed by Hautus (1995) and Brown and White (2005). We added 0.5 to the number of hits and the number of false alarms and 1 to the number of signal trials and the number of noise trials. Two two-way analyses of variance were performed to compare the memory scores of the new/old recognition task between our four experimental groups depending on the level of immersion and realism of the presentation: one for the data collected during the post-test (immediately after the intervention) and one for the data collected during the follow-up test (one week after the intervention). We report analyses of covariance (with the covariates of age, gender, theory of mind, and language abilities) in Appendix D. Adding the covariates did not change the results reported here in terms of rejecting the null hypotheses. We depict all the data in Figure 4 below.

Post-Test (Immediately After the Intervention). On the first day, we had to exclude three participants because they were not able to finish the task due to a lack of time and three other participants due to technical problems. In the end, we carried out the analysis with 158 children, including 34 in the realistic D-VR condition, 46 in the realistic IVR condition, 37 in the anthropomorphic D-VR condition, and 41 in the anthropomorphic IVR condition.

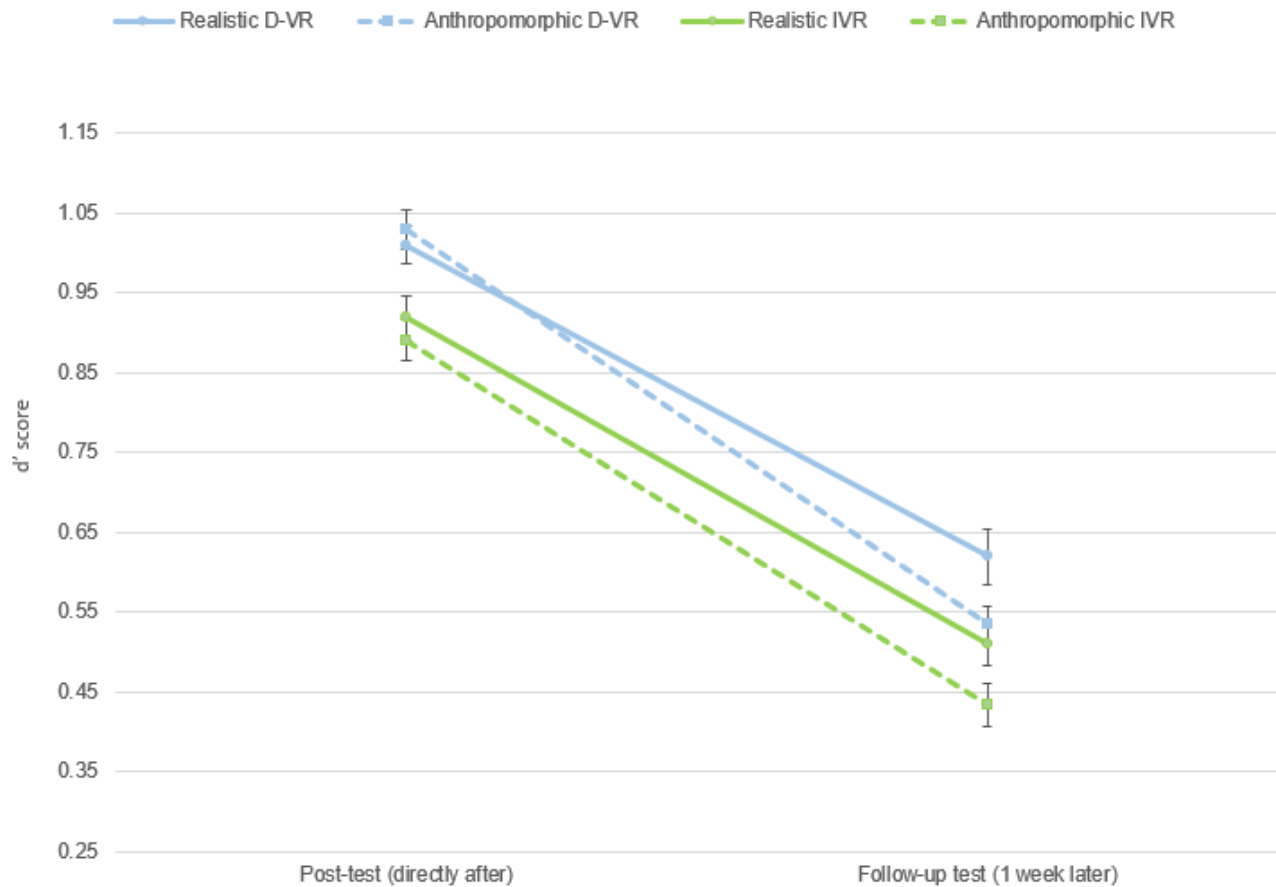
We observed a significant impact of immersion on the results for the new/old recognition task ($F(1, 154) = 18.331$; $p < .001$, $\eta_p^2 = .106$). Contrary to our expectations, participants in the D-VR condition ($M = 1.02$, $SD = .143$) performed significantly better than participants in the IVR condition ($M = .906$, $SD = .180$). On the contrary, we did not observe a significant difference between the realistic ($M = .956$, $SD = .175$) and anthropomorphic ($M = .957$, $SD = .173$) conditions for the results of the new/old recognition task ($F(1, 154) = .008$; $p = .927$, $\eta_p^2 < .001$). Further, we found no significant interaction between the levels of realism and immersion of the presentation ($F(1, 154) = .936$; $p = .335$, $\eta_p^2 = .006$).

Follow-Up Test (One Week After the Intervention). On the second day, we had to exclude seven participants because they dropped out or were not able to finish the testing, 10 participants due to technical problems, and one participant for a score lower than three times the standard deviation. In the end, we carried out the analysis on 146 children, including 32 in the realistic D-VR condition, 44 in the realistic IVR condition, 39 in the anthropomorphic D-VR condition, and 31 in the anthropomorphic IVR condition.

On the second day, we found, once again, a significant effect for the level of immersion on the children's performance on the new/old recognition task ($F(1, 142) = 13.389$; $p < .001$, $\eta_p^2 = .086$). Again, participants performed better in the D-VR condition ($M = .573$, $SD = .177$) than in the IVR condition ($M = .478$, $SD = .175$). Contrary to the first day, on the second day, one week later, we observed a significant difference in the children's performance on the new/old recognition task depending on the level of realism of the presentation ($F(1, 142) = 7.722$; $p = .006$, $\eta_p^2 = .052$). Participants performed better in the realistic condition ($M = .556$, $SD = .196$) than in the anthropomorphic condition ($M = .490$, $SD = .160$). Moreover, we found no significant interaction between the levels of immersion and realism of the presentation ($F(1, 142) = .019$; $p = .890$, $\eta_p^2 < .001$).

Figure 4

d' Scores for the New/Old Recognition Task



Note. These scores were calculated according to the realism of the presentation (realistic/anthropomorphic) and the level of immersion (IVR/D-VR) during the post-test directly after the presentation and during the follow-up test one week later. Error bars represent one SEM.

Quiz Task

The children were able to obtain a maximum of 12 points on the quiz task. The higher the score, the better the children performed. Two two-way analyses of variance were performed to compare the memory scores of the quiz task between our four experimental groups depending on the level of immersion and realism of the presentation, including one for the data collected during the post-test (immediately after the intervention) and one for the data collected during the follow-up test (one week after the intervention). We report the analyses of covariance (with the covariates of age, gender, theory of mind, and language abilities) in Appendix E. Adding the covariates did not change the results reported here in terms of rejecting the null hypotheses.

Post-Test (Immediately After the Intervention). On the first day, we had to exclude one participant because of the inability to finish the task due to lack of time and two other participants for a score lower than three times the

standard deviation. In the end, we carried out the analysis on 161 children, including 34 in the realistic D-VR condition, 45 in the realistic IVR condition, 39 in the anthropomorphic D-VR condition, and 43 in the anthropomorphic IVR condition.

As with the new/old recognition task, we observed a significant impact of immersion on the results of the quiz task for the first day ($F(1, 157) = 8.792$; $p = .003$, $\eta_p^2 = .053$). Once again, participants in the D-VR condition ($M = 6.34$, $SD = 1.25$) performed significantly better than participants in the IVR condition ($M = 5.67$, $SD = 1.52$). We did not observe a significant difference between the realistic ($M = 5.92$, $SD = 1.56$) and anthropomorphic ($M = 6.02$, $SD = 1.32$) conditions in the results of this task ($F(1, 157) = .204$; $p = .652$, $\eta_p^2 = .001$). We found no significant interaction between the levels of realism and immersion in the presentation ($F(1, 157) = 2.069$; $p = .152$, $\eta_p^2 = .013$).

Follow-Up Test (One Week After the Intervention). On the second day, we had to exclude nine participants because they dropped out or were not able to finish the testing, and one participant for a score lower than three times the standard deviation. In the end, we carried out the analysis on 154 children, including 32 in the realistic D-VR condition, 45 in the realistic IVR condition, 39 in the anthropomorphic D-VR condition, and 38 in the anthropomorphic IVR condition.

Contrary to the previous day and the new old/recognition task, we did not find a difference in the quiz scores depending on the level of immersion ($F(1, 150) = 2.95$; $p = .088$, $\eta_p^2 = .019$). Participants in the D-VR conditions ($M = 6.24$, $SD = 1.20$) performed as well as participants in the IVR conditions ($M = 5.82$, $SD = 1.59$). We also found no difference in the quiz results between participants in the realistic ($M = 5.88$, $SD = 1.50$) and anthropomorphic ($M = 6.14$, $SD = 1.35$) conditions on the second day ($F(1, 150) = 1.20$; $p = .275$, $\eta_p^2 = .008$). Further, we found no significant interaction between the levels of realism and immersion of the presentation ($F(1, 150) = 2.39$; $p = .124$, $\eta_p^2 = .016$).

Transfer

The children were able to obtain a maximum of four points (one point for each item correctly remembered as Chinese food) on the transfer task. The higher the score, the better the children performed. The children were tested only on transfer in the follow-up testing phase. Once again, a two-way analysis of variance was performed to compare the transfer scores between our four experimental groups, depending on the levels of immersion and realism of the presentation. We had to exclude nine participants because they dropped out or were not able to finish the testing, and one participant for a score lower than three times the standard deviation. In the end, we carried out the analysis on 154 children, including 33 in the realistic D-VR condition, 45 in the realistic IVR condition, 38 in the anthropomorphic D-VR

condition, and 38 in the anthropomorphic IVR condition. We report the analyses of covariance (with the covariates of age, gender, theory of mind, and language abilities) in Appendix F. Adding the covariates did not change the results reported here in terms of rejecting the null hypotheses.

As for the other tasks, we also observed a significant impact of the immersion level on the transfer task ($F(1, 150) = 5.286$; $p = .023$, $\eta_p^2 = .034$). Again, participants in the D-VR condition ($M = 3.23$, $SD = .778$) performed significantly better than participants in the IVR condition ($M = 2.94$, $SD = .817$). We did not observe a significant difference between the realistic ($M = 3.14$, $SD = .768$) and anthropomorphic ($M = 3.00$, $SD = .849$) conditions ($F(1, 150) = 1.651$; $p = .201$, $\eta_p^2 = .011$). We also found no significant interaction between the levels of realism and immersion ($F(1, 150) = .073$; $p = .787$, $\eta_p^2 < .001$).

Perception of the Avatar. Three questions were asked to the children concerning their perception of the avatar and the truthfulness of its words. As a reminder, the child had the choice between two propositions for each of these questions. Although the children were encouraged to answer with one of the two propositions, some of them failed to provide an answer and were therefore not taken into account in the analyses. Each question was analyzed individually with respect to the realism of the presentation and the medium used. To do this, we computed binomial logistic regressions for each of the three questions, with immersion (D-VR/IVR) and realism (realistic/non-realistic) as factors. The count of the children's responses to each question is reported in Appendix I.

For the first question ("What did the character in the video look like?"), we had to exclude 18 participants because they were not able to select an answer from among the propositions. In the end, we carried out the analysis for 146 children, including 28 in the realistic D-VR condition, 43 in the realistic IVR condition, 34 in the anthropomorphic D-VR condition, and 37 in the anthropomorphic IVR condition. The logistic regression model was statistically significant ($\chi^2(2) = 169$, $p < .001$). The realism of the presentation was a significant predictor of how children perceived the avatar ($Z = 6.108$, $p < .001$). The odds of classifying the avatar as a little girl were 2626.210 (95% CI 209.920–32855.180) in the realistic condition. Immersion did not significantly predict how the child perceived the avatar ($Z = .648$, $p = .517$).

For the second question ("Do you believe what the character told you? Did the character tell the truth?"), we had to exclude 23 participants because they failed to provide an answer. In the end, we carried out the analysis on 141 children, including 30 in the realistic D-VR condition, 38 in the realistic IVR condition, 33 in the anthropomorphic D-VR condition, and 40 in the anthropomorphic IVR condition. The logistic regression model was not statistically significant

($\chi^2(2) = .379$, $p = .828$). The count of the children's responses reported in Appendix H illustrates that the children overall believed what the character told them.

For the third question ("Can a character like this really exist in our world? Could we meet a character like this in our world?"), we had to exclude 16 participants because they failed to provide an answer. In the end, we carried out the analysis on 146 children, including 29 in the realistic D-VR condition, 41 in the realistic IVR condition, 40 in the anthropomorphic D-VR condition, and 36 in the anthropomorphic IVR condition. The logistic regression model was statistically significant ($\chi^2(2) = 9.61$, $p = .008$). The realism of the presentation was a significant predictor of how the children perceived the possible existence of the avatar ($Z = 2.75$, $p = .006$). The odds of classifying the avatar as possibly existing in the real world were 2.831 (95% CI 1.3482–5.945) in the realistic condition. Immersion was not a significant predictor of responses to this question ($Z = 1.16$, $p = .244$).

Enjoyment and Perceived Purpose of the Virtual Presentation. To assess the enjoyment and perceived purpose of the virtual presentation, the children had to answer two questions afterward. First, the children were asked whether they enjoyed the presentation on a five-point Likert scale. To investigate the impact of realism and immersion on their appreciation of the presentation, we computed a two-way analysis of variance. We had to exclude one participant because of the inability to select an answer from among the propositions. We did not observe a significant difference between the IVR ($M = 4.70$, $SD = .714$) and D-VR ($M = 4.69$, $SD = .701$) conditions on the appreciation of the presentation ($F(1, 159) = .009$; $p = .925$, $\eta_p^2 = .000$). We also did not observe a significant difference between the realistic ($M = 4.67$, $SD = .758$) and anthropomorphic ($M = 4.72$, $SD = .653$) conditions with respect to their effect on the appreciation of the presentation ($F(1, 159) = .273$; $p = .602$, $\eta_p^2 = .002$). We found no significant interaction between the levels of realism and immersion in the presentation ($F(1, 159) = .254$; $p = .615$, $\eta_p^2 = .002$).

Second, the children were asked if they thought we had showed them the presentation for fun or for learning. We had to exclude five participants because they were unable to select an answer from among the propositions. In the end, we carried out the analysis on 159 children, including 33 in the realistic D-VR condition, 46 in the realistic IVR condition, 38 in the anthropomorphic D-VR condition, and 42 in the anthropomorphic IVR condition. To investigate the impact of realism and immersion, we computed a binomial logistic regression with immersion (D-VR/IVR) and realism (realistic/anthropomorphic) as factors. The logistic regression model was not statistically significant ($\chi^2(2) = 0.077$, $p = 0.962$). The count of the children's responses reported in Appendix J illustrates that the children considered the overall presentation to be for learning.

Discussion

In the present study, we investigated the impact of fantasy and immersion on young children's recall of cultural information. The findings provide evidence that young children are sensitive to the appearance of the avatar doing the presentation; specifically, whether the avatar was realistic or anthropomorphic had an impact on their memory performance one week after the intervention. Contrary to our expectations, children in the D-VR conditions outperformed children working with headsets (IVR conditions). Gender, age, theory of mind, and verbal abilities did not have an impact on the significance of our results. From a more practical point of view, almost all the children very much appreciated the presentation, regardless of its realistic and immersive nature (mean of 4.69 on a 5-point Likert scale); they also considered the presentation to have been delivered for learning. In the following sections, we provide possible explanations of our findings.

Children Learning with Fantasy

As expected, the children performed better in the realistic condition than in the anthropomorphic condition. However, this difference was significant only one week after the virtual presentation and only for the new/old recognition task. It seems that at first, children memorized the information equally well, independently of the realism of the avatar. These results are consistent with the literature indicating that children can learn from fiction (Aydin et al., 2021; Conrad et al., 2021), although unlike some other authors (e.g., Hopkins & Lillard, 2021), we did not find that fantastical elements significantly enhanced learning. On the contrary, our results for the new/old recognition task one week after the virtual presentation suggest an advantage of realistic information. It is possible that the information that children consider realistic is better consolidated in their long-term memory. This finding is consistent with the results of Bonus and Mares (2019), who found that children remembered and understood a cultural presentation better when they judged it to be realistic, but only in the long-term period of 5 to 12 days after the intervention.

Why should it be easier for young children to recall the elements of a realistic presentation? First, when presented with realistic information, children can import more real information and thus create a solid mental representation of the information by using prior knowledge of the world to support their understanding. Furthermore, information from an anthropomorphic character might interfere with memory processes; fantastical material might enhance the cognitive demand of the task and hamper the encoding/retrieval process (Fisch, 2000; Fisch et al., 2005). Yet another explanation for the better memory of information from a realistic avatar is that children might invest more effort in processing realistic material since it is more relevant to the real world (Mares & Sivakumar, 2014). It remains an

open question whether children put more effort into processing elements from a realistic avatar, or whether elements presented by a realistic avatar are simply easier to recall in a long-term old/new recognition task.

We believe that one of the possible reasons why we did not find an effect of realism with the quiz task is because we used the exact same questions immediately after the intervention and one week after the presentation. Therefore, on the second quiz, one week after the presentation, the children could rely not only on what they learned from the presentation but also on the quiz itself, which might have reinforced and decontextualized the learned cultural information. However, it is surprising that we did not find an impact of realism on the transfer task, which was used only one week later. A possible explanation here is that the task had too few items to detect the effect. Although the effect is not generalizable among our measurements, we think that the result is interesting, and that it deserves further research. Future studies should investigate the conditions under which realism has an impact on learning. In their recent literature review, Weisberg and Richert (2022) propose that the quality of learning with material containing fantastical elements is dependent on multiple factors (i.e., degree of realism, target educational content, and depth of integration of the fantastical elements in the story). Future studies should also test the extent to which these different factors are important for short- and long-term retention.

Finally, the children correctly differentiated the two avatars that were presented to them (young girl vs. anthropomorphic animal). One of the elements that we were able to identify with our question as allowing them to distinguish the two avatars as more or less realistic was the possibility that the avatar exists in the real world, and that one could meet it. The children in the realistic condition thought significantly more often that one could meet the avatar of the young girl in the real world. On the contrary, children in anthropomorphic conditions thought significantly more often that the anthropomorphic avatar did not exist in the real world and could not be met anywhere on Earth. We did not observe any other significant differences between the conditions in the responses to our exploratory questions. Future research should continue to investigate the different characteristics that make an avatar more or less realistic and the impact of such characteristics on learning.

Children Learning with Immersive Virtual Reality

Contrary to our hypothesis, children performed significantly worse in the IVR conditions across tasks (new/old recognition task, quiz task, and transfer task). Based on previous research (Babu et al., 2018; Makransky & Mayer, 2022), we expected that children would retrieve information at least as well in the IVR as in the D-VR conditions. A first possible explanation for this finding is that IVR induced too much extraneous processing compared to the generative

processing that was thought to result from immersion. This interpretation is in line with the findings of Makransky et al. (2019), who reported that IVR elicited a stronger feeling of presence and more liking but less learning.

What were the possible drivers of extraneous processing in our immersive environment? First, the novelty of IVR could have generated increased cognitive processing; in other words, children in the D-VR conditions may have performed better because they were more familiar with the learning medium when compared to children working with IVR. An investigation by Bernath et al. (2020) showed that in Switzerland, out of nearly 900 households with children aged four to six, 79% were equipped with at least one tablet. VR headsets are much less common, especially among young children. In our study, 54% of the households specified that their children used a tablet at least once a month, while it was only 11.3% of the children who had already tried a VR headset, and 8.3% only on one occasion. According to the theory of multimedia learning (Mayer, 2009), habituation to the use of a VR headset can be considered extraneous processing and would offset the generative processing linked to the benefits of immersion offered by IVR. Even though the children in our study had a short phase of familiarization with the virtual environment and the corresponding medium before the virtual presentation, this might not have been sufficient to acquire the same level of familiarization between headsets and tablets. It is plausible that children might profit more from IVR if they have a longer phase of habituation or use IVR repeatedly. This is what Meyer et al. (2019) observed when they added IVR pre-training phases to their IVR teaching design.

Another driver of extraneous processing in our environment might have been the details of the graphics, especially assets that were not directly useful for learning purposes. Despite our careful design of the environment, it is still possible that the level of graphic detail in the VR environment was too high and created extraneous processing. These details would have been particularly disturbing in the IVR conditions, where the level of resolution was higher, and the environment fully surrounded the children, as compared to the lower-resolution D-VR conditions, where the children were sitting in front of the tablet (i.e., a more restricted field of view). It has been suggested that to fully benefit from IVR, the virtual environment should not be too detailed, especially regarding elements that are not related to learning. Nevertheless, this reduction in detail should not be at the expense of a loss of immersion or of the learner's ability to contextualize the learning materials (Makransky & Mayer, 2022).

Yet another explanation for the D-VR advantage we found is the concordance of the medium used during the encoding and recall tasks. Indeed, while the encoding was done on a tablet or with a VR headset, depending on the condition, the new/old recognition task was always carried out on a computer, a medium closer to the tablet than to

the VR headset. In this regard, children in the D-VR conditions may have had an advantage over children in the IVR conditions. Future research should carefully compare the use of different media (IVR, D-VR, and computer experiments) during the encoding and recall phases.

Conclusion

In view of our findings and the rest of the literature, we continue to note mixed results regarding the effectiveness of IVR and materials containing fantastical elements in learning experiences. Regarding VR, our results clearly point to a benefit of D-VR when compared to IVR for four-to-six-year-olds. As discussed above, follow-up studies (e.g., manipulating the degree of habituation with the medium, the level of detail of the VR environment, or using IVR to assess memory performance) are needed to fully understand this finding. To our knowledge, this is one of the first studies where such young children had to learn with a VR headset, and before proposing the use of IVR with preschool children, much more research is needed about children's learning in VR. This type of research is needed to assess the full educational potential of these methods and to avoid potential detrimental effects, as it is currently unknown whether VR adds value to more traditional materials in preschool settings.

Finally, our results seem to suggest that children do not retrieve information from an anthropomorphic avatar as well as they retrieve information from a realistic avatar when considering memory consolidation (i.e., assessment one week after the presentation). It is important to specify that this study has looked at the impact of an anthropomorphic avatar on the recall of cultural information, and that because fantasy appears to place more pressure on cognitive processes, it may have some value in supporting cognitive development in other contexts, such as fantastical play (Thibodeau et al., 2016). As these results were significant only for the new/old recognition task, it is important to continue investigating the impact of realism on children's learning and see if these results can be replicated and generalized to other tasks.

It is also important to note that the length of the presentation in the study (5 minutes) is below the standard length of TV programs aimed at children in this age group. The shorter duration of the presentation was chosen mainly due to limited scientific data on the effects of IVR on such a young population. In prioritizing ethical considerations, we decided not to make the presentation too long, to avoid any potential negative effects. Replicating this study with a longer duration in IVR vs. D-VR would be interesting, as it would better align with children's everyday experience with educational TV programs.

To summarize, this study provides evidence that the type of medium (IVR vs. D-VR) and the type of material (anthropomorphic vs. realistic) influence the way information is processed, encoded, and memorized by young children. Even though the underlying mechanisms remain unclear, these findings highlight the importance of carefully selecting and designing educational material for young children when seeking to support their recall of cultural information.

References

- Araiza-Alba, P., Keane, T., Chen, W. S., & Kaufman, J. (2021). Immersive virtual reality as a tool to learn problem-solving skills. *Computers & Education*, 164, 104121. <https://doi.org/10.1016/j.compedu.2020.104121>
- Aydin, E., Ilgaz, H., & Allen, J. W. P. (2021). Preschoolers' learning of information from fantastical narrative versus expository books. *Journal of Experimental Child Psychology*, 209, 105170. <https://doi.org/10.1016/j.jecp.2021.105170>
- Babu, S., Krishna, S., & Bhavani, R. R. (2018). Virtual reality learning environments for vocational education: A comparison study with conventional instructional media on knowledge retention. *2018 IEEE 18th International Conference on Advanced Learning Technologies (ICALT)*, 385–389. <https://doi.org/10.1109/ICALT.2018.00094>
- Bernath, J., Waller, G., & Meidert, U. (2020). *ADELE+: Der Medienumgang von Kindern im Vorschulalter (4-6 Jahre)—Chancen und Risiken für die Gesundheit*.
- Blascovich, J., Loomis, J., Beall, A. C., Swinth, K. R., Hoyt, C. L., & Bailenson, J. N. (2002). Immersive virtual environment technology as a methodological tool for social psychology. *Psychological Inquiry*, 13(2), 103–124.
- Bonus, J. A., & Mares, M.-L. (2018). When the Sun Sings Science, Are Children Left in the Dark? Representations of Science in Children's Television and Their Effects on Children's Learning. *Human Communication Research*, 44(4), 449-472. <https://doi.org/10.1093/hcr/hqy009>
- Bonus, J. A., & Mares, M.-L. (2019). Learned and remembered but rejected: Preschoolers' reality judgments and transfer from *Sesame Street*. *Communication Research*, 46(3), 375–400. <https://doi.org/10.1177/0093650215609980>
- Borzekowski, D. L. G., & Macha, J. E. (2010). The role of Kilimani Sesame in the healthy development of Tanzanian preschool children. *Journal of Applied Developmental Psychology*, 31(4), 298-305. <https://doi.org/10.1016/j.appdev.2010.05.002>

- Buttussi, F., & Chittaro, L. (2018). Effects of different types of virtual reality display on presence and learning in a safety training scenario. *IEEE Transactions on Visualization and Computer Graphics*, 24(2), 1063–1076.
<https://doi.org/10.1109/TVCG.2017.2653117>
- Chlebuch, N., Bodas, A., & Weisberg, D. S. (2022). What does the Cat in the Hat know about that? An analysis of the educational and unrealistic content of children’s narrative science media. *Psychology of Popular Media*.
<https://doi.org/10.1037/ppm0000388>
- Clements, W. A., & Perner, J. (1994). Implicit understanding of belief. *Cognitive Development*, 9(4), 377–395.
[https://doi.org/10.1016/0885-2014\(94\)90012-4](https://doi.org/10.1016/0885-2014(94)90012-4)
- Cole, C. F., Arafat, C., Tidhar, C., Tafesh, W. Z., Fox, N. A., Killen, M., Ardila-Rey, A., Leavitt, L. A., Lesser, G., Richman, B. A., & Yung, F. (2003). The educational impact of Rechov Sumsum/Shara’a Simsim : A Sesame Street television series to promote respect and understanding among children living in Israel, the West Bank, and Gaza. *International Journal of Behavioral Development*, 27(5), 409-422. <https://doi.org/10.1080/01650250344000019>
- Conrad, M., Marcovitch, S., & Boseovski, J. J. (2021). The friendly fossa: The effect of anthropomorphic language on learning about unfamiliar animals through both storybooks and live animal experiences. *Journal of Experimental Child Psychology*, 201, 104985. <https://doi.org/10.1016/j.jecp.2020.104985>
- Corriveau, K. H., & Harris, P. L. (2015). Children’s developing realization that some stories are true: Links to the understanding of beliefs and signs. *Cognitive Development*, 34, 76–87.
<https://doi.org/10.1016/j.cogdev.2014.12.005>
- Fisch, S. M. (2000). A capacity model of children’s comprehension of educational content on television. *Media Psychology*, 2(1), 63–91. https://doi.org/10.1207/S1532785XMEP0201_4
- Fisch, S. M., Kirkorian, H., & Anderson, D. (s. d.). *Transfer of learning in informal education*.
- Ganea, P. A., Canfield, C. F., Simons-Ghafari, K., & Chou, T. (2014). Do cavies talk? The effect of anthropomorphic picture books on children’s knowledge about animals. *Frontiers in Psychology*, 5.
<https://doi.org/10.3389/fpsyg.2014.00283>
- Geerdts, M. S. (2016). (Un)real animals: Anthropomorphism and early learning about animals. *Child Development Perspectives*, 10(1), 10–14. <https://doi.org/10.1111/cdep.12153>

- Geerds, M. S., Van de Walle, G. A., & LoBue, V. (2016). Learning About Real Animals From Anthropomorphic Media. *Imagination, Cognition and Personality*, 36(1), 5-26. <https://doi.org/10.1177/0276236615611798>
- Goldstein, T. R., & Alperson, K. (2020). Dancing bears and talking toasters: A content analysis of supernatural elements in children's media. *Psychology of Popular Media*, 9(2), 214-223. <https://doi.org/10.1037/ppm0000222>
- Goldstein, T. R., Thompson, B. N., & Kanumuru, P. (2022). Do embodiment and fictionality affect young children's learning? *Journal of Experimental Child Psychology*, 213, 105275. <https://doi.org/10.1016/j.jecp.2021.105275>
- Hopkins, E. J., & Lillard, A. S. (2021). The Magic School Bus dilemma: How fantasy affects children's learning from stories. *Journal of Experimental Child Psychology*, 210, 105212. <https://doi.org/10.1016/j.jecp.2021.105212>
- Hopkins, E. J., & Weisberg, D. S. (2021). Investigating the effectiveness of fantasy stories for teaching scientific principles. *Journal of Experimental Child Psychology*, 203, 105047. <https://doi.org/10.1016/j.jecp.2020.105047>
- Larsen, N. E., Lee, K., & Ganea, P. A. (2018). Do storybooks with anthropomorphized animal characters promote prosocial behaviors in young children? *Developmental Science*, 21(3), e12590. <https://doi.org/10.1111/desc.12590>
- Leder, J., Horlitz, T., Puschmann, P., Wittstock, V., & Schütz, A. (2019). Comparing immersive virtual reality and PowerPoint as methods for delivering safety training: Impacts on risk perception, learning, and decision making. *Safety Science*, 111, 271-286. <https://doi.org/10.1016/j.ssci.2018.07.021>
- Li, H., Eisen, S., & Lillard, A. S. (2019). Anthropomorphic media exposure and preschoolers' anthropomorphic thinking in China. *Journal of Children and Media*, 13(2), 149-162. <https://doi.org/10.1080/17482798.2019.1570960>
- Makransky, G., & Lilleholt, L. (2018). A structural equation modeling investigation of the emotional value of immersive virtual reality in education. *Educational Technology Research and Development*, 66(5), 1141-1164. <https://doi.org/10.1007/s11423-018-9581-2>
- Makransky, G., & Mayer, R. E. (2022). Benefits of taking a virtual field trip in immersive virtual reality: Evidence for the immersion principle in multimedia learning. *Educational Psychology Review*, 34(3), 1771-1798. <https://doi.org/10.1007/s10648-022-09675-4>
- Makransky, G., Terkildsen, T. S., & Mayer, R. E. (2019). Adding immersive virtual reality to a science lab simulation causes more presence but less learning. *Learning and Instruction*, 60, 225-236. <https://doi.org/10.1016/j.learninstruc.2017.12.007>

- Mares, M.-L., & Sivakumar, G. (2014). “Vámonos means go, but that’s made up for the show”: Reality confusions and learning from educational TV. *Developmental Psychology*, 50(11), 2498–2511. <https://doi.org/10.1037/a0038041>
- Martarelli, C. S., & Mast, F. W. (2013). Is it real or is it fiction? Children’s bias toward reality. *Journal of Cognition and Development*, 14(1), 141–153. <https://doi.org/10.1080/15248372.2011.638685>
- Martarelli, C. S., Mast, F. W., Läge, D., & Roebbers, C. M. (2015). The distinction between real and fictional worlds: Investigating individual differences in fantasy understanding. *Cognitive Development*, 36, 111–126. <https://doi.org/10.1016/j.cogdev.2015.10.001>
- Mayer, R. E. (2009). *Multimedia learning* (2nd ed.). Cambridge University Press. <https://doi.org/10.1017/CBO9780511811678>
- Mayr, S., & Buchner, A. (2006). Evidence for episodic retrieval of inadequate prime responses in auditory negative priming. *Journal of Experimental Psychology: Human Perception and Performance*, 32(4), 932–943. <https://doi.org/10.1037/0096-1523.32.4.932>
- Meyer, O. A., Omdahl, M. K., & Makransky, G. (2019). Investigating the effect of pre-training when learning through immersive virtual reality and video: A media and methods experiment. *Computers & Education*, 140, 103603. <https://doi.org/10.1016/j.compedu.2019.103603>
- Nguyentran, G., & Weisberg, D. S. (2023). Should the cat in the hat keep talking like that? Educational correlates of anthropomorphism in children’s science media. *Psychology of Popular Media*. <https://doi.org/10.1037/ppm0000467>
- Olmos-Raya, E., Ferreira-Cavalcanti, J., Contero, M., Castellanos, M. C., Giglioli, I. A. C., & Alcañiz, M. (2018). Mobile virtual reality as an educational platform: A pilot study on the impact of immersion and positive emotion induction in the learning process. *EURASIA Journal of Mathematics, Science and Technology Education*, 14(6). <https://doi.org/10.29333/ejmste/85874>
- Parong, J., & Mayer, R. E. (2018). Learning science in immersive virtual reality. *Journal of Educational Psychology*, 110(6), 785–797. <https://doi.org/10.1037/edu0000241>
- Perner, J., Leekam, S. R., & Wimmer, H. (1987). Three-year-olds’ difficulty with false belief: The case for a conceptual deficit. *British Journal of Developmental Psychology*, 5(2), 125–137. <https://doi.org/10.1111/j.2044-835X.1987.tb01048.x>

- Richert, R. A., & Schlesinger, M. A. (2017). The role of fantasy-reality distinctions in preschoolers' learning from educational video : Fantasy-reality distinctions predict preschoolers' learning from video. *Infant and Child Development*, 26(4), e2009. <https://doi.org/10.1002/icd.2009>
- Richert, R. A., & Schlesinger, M. A. (2022). Relations between fantasy and transfer of learning from storybooks. *Journal of Experimental Child Psychology*, 222, 105474. <https://doi.org/10.1016/j.jecp.2022.105474>
- Richert, R. A., & Smith, E. I. (2011). Preschoolers' quarantining of fantasy stories. *Child Development*, 82(4), 1106–1119. <https://doi.org/10.1111/j.1467-8624.2011.01603.x>
- Rupp, M. A., Odette, K. L., Kozachuk, J., Michaelis, J. R., Smither, J. A., & McConnell, D. S. (2019). Investigating learning outcomes and subjective experiences in 360-degree videos. *Computers & Education*, 128, 256–268. <https://doi.org/10.1016/j.compedu.2018.09.015>
- Sharon, T., & Woolley, J. D. (2004). Do monsters dream? Young children's understanding of the fantasy/reality distinction. *British Journal of Developmental Psychology*, 22(2), 293–310. <https://doi.org/10.1348/026151004323044627>
- Slater, M., & Wilbur, S. (1997). A framework for immersive virtual environments (FIVE): Speculations on the role of presence in virtual environments. *Presence: Teleoperators and Virtual Environments*, 6(6), 603–616. <https://doi.org/10.1162/pres.1997.6.6.603>
- Slater, M. (2003). A note on presence terminology. *Presence Connect*, 3(3), 1–5.
- Smith, S. M., Glenberg, A., & Bjork, R. A. (1978). Environmental context and human memory. *Memory & Cognition*, 6(4), 342–353. <https://doi.org/10.3758/BF03197465>
- Smith, S. A. (2019). Virtual reality in episodic memory research: A review. *Psychonomic Bulletin & Review*, 26(4), 1213–1237. <https://doi.org/10.3758/s13423-019-01605-w>
- Walker, C. M., Gopnik, A., & Ganea, P. A. (2015). Learning to learn from stories: Children's developing sensitivity to the causal structure of fictional worlds. *Child Development*, 86(1), 310–318. <https://doi.org/10.1111/cdev.12287>
- Weisberg, D. S., & Hopkins, E. J. (2020). Preschoolers' extension and export of information from realistic and fantastical stories. *Infant and Child Development*, 29(4). <https://doi.org/10.1002/icd.2182>
- Weisberg, D. S., & Richert, R. A. (2022). How, when, and what do young children learn from fictional stories? *Journal of Experimental Child Psychology*, 221, 105445. <https://doi.org/10.1016/j.jecp.2022.105445>

Wechsler, D. (2012). *WPPSI-IV: Wechsler preschool and primary scale of intelligence* (4th ed.). San Antonio, TX: Psychological Corporation.

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, 13(1), 103–128. [https://doi.org/10.1016/0010-0277\(83\)90004-5](https://doi.org/10.1016/0010-0277(83)90004-5)

Woolley, J. D., & Cox, V. (2007). Development of beliefs about storybook reality. *Developmental Science*, 10(5), 681–693. <https://doi.org/10.1111/j.1467-7687.2007.00612.x>

Appendix A

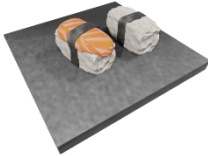
Visual stimuli used for the new/old recognition task. The 12 “old” stimuli were shown to the children during the virtual reality presentation, while the 12 “new” stimuli were not. Among the 12 “old” stimuli, six were 3D models of dishes or objects related to cooking, three were landscapes of places to visit in China, and three were images of activities to do in China). Among the 12 “new” stimuli, six were 3D models of dishes or objects related to cooking, three were landscapes of places to visit in India, and three were images of activities to do in India. The 24 visual stimuli were separated into two different sets of 12 stimuli so that the children were presented with different items for the post-test and the follow-up test.

Set A

Old stimuli	New stimuli
 Lotus root	 Empanadas
 Dragon fruit	 Cabbage
 Bao	 Cake

	
Tai Chi	Indian landscape
	
Kite	Taj Mahal
	
Huang Shan	Polo

Set B

Old stimuli	New stimuli
	
Durian	Sushi



Chopsticks



Broth



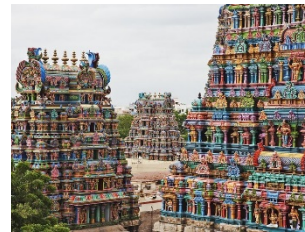
Zongzi



Kiwano



Great Wall of China



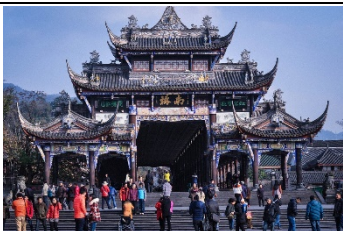
Tamil Nadu



Karaoke



Dance



Temple



Boat

Appendix B

Questions asked of the children for the quiz task. Each question was accompanied by four suggested answers.

When the question was visual, larger versions of the corresponding proposed images were shown to the children.

1. What is the name of the character who talks to you in the story?

A. Mei B. Liu C. Lin-Lin D. Ming

2. What country does this character come from?

A. Japan B. China C. India D. Korea

3. What does her first name mean?

A. Tree B. Moon C. Leaf D. Forest

4. What color are her clothes?

A. Red B. Blue C. Green D. Purple

5. How do you say hello in Chinese?

A. Nihao B. Ani C. Onia D. Boni

6. How do you write her first name in Chinese?

A. 林林 B. 字字 C. 含含 D. 品品

7. Which of these flags represents China?

A.



B.



C.



D.



8. Where is China on the map?

A.



B.



C.



D.



Appendix C

Image presented to the children during the transfer task. In the picture, there are 10 different food items, among which the child was asked to select four items related to China. Among these 10 items, five are variations of the 3D models used in the virtual presentation, one is a variation of a “new” 3D model from the new/old recognition task, and four are 2D images from the internet of non-typical Chinese food.



Appendix D

Analysis of covariance for the new/old recognition task during the post-test, with realism and immersion as independent variables, and gender, age, theory of mind (ToM), and verbal abilities as covariates.

	$F(1,143)$	p	η_p^2
Gender	0.765	0.383	0.005
Age	1.534	0.218	0.011
ToM abilities	2.619	0.108	0.018
Verbal abilities	5.499	0.020	0.037
Immersion	9.107	0.003	0.060
Realism	< 0.001	0.983	< 0.001
Immersion * realism	0.003	0.701	0.001

Analysis of covariance for the new/old recognition task during the follow-up test, with realism and immersion as independent variables, and gender, age, theory of mind, and verbal abilities as covariates.

	$F(1,136)$	P	η_p^2
Gender	0.016	0.900	< 0.001
Age	3.763	0.054	0.027
ToM abilities	2.504	0.116	0.018
Verbal abilities	0.916	0.340	0.007
Immersion	10.928	0.001	0.074
Realism	7.574	0.007	0.053
Immersion * realism	0.002	0.962	< 0.001

Appendix E

Analysis of covariance for the quiz task during the post-test, with realism and immersion as independent variables, and gender, age, theory of mind, and verbal abilities as covariates.

	$F(1,145)$	P	η_p^2
Gender	1.489	0.224	0.010
Age	3.221	0.075	0.022
ToM abilities	5.083	0.026	0.034
Verbal abilities	12.764	< 0.001	0.081
Immersion	4.190	0.042	0.028
Realism	0.345	0.558	0.002
Immersion * realism	0.629	0.429	0.004

Analysis of covariance for the quiz task during the follow-up test, with realism and immersion as independent variables, and gender, age, theory of mind, and verbal abilities as covariates.

	$F(1,146)$	P	η_p^2
Gender	0.610	0.436	0.004
Age	3.406	0.067	0.023
ToM abilities	4.298	0.040	0.029
Verbal abilities	2.411	0.123	0.016
Immersion	1.303	0.255	0.009
Realism	1.969	0.163	0.013
Immersion * realism	2.241	0.137	0.015

Appendix F

Analysis of covariance for the transfer task during the follow-up test, with realism and immersion as independent variables, and gender, age, theory of mind, and verbal abilities as covariates.

	<i>F</i> (1,146)	<i>P</i>	η_p^2
Gender	0.546	0.461	0.004
Age	0.279	0.598	0.002
ToM abilities	2.857	0.093	0.019
Verbal abilities	6.526	0.012	0.043
Immersion	2.668	0.105	0.018
Realism	1.862	0.174	0.013
Immersion * realism	0.601	0.439	0.004

Appendix G

We computed a 2 x 2 x 2 ANOVA for mixed designs. The level of immersion and realism of the presentation served as between-subjects variables, while the time of testing (post-test/follow-up test) was considered as within-subjects variable. We had to exclude 10 participants because they dropped out or were not able to finish the testing, 11 participants due to technical problems, and one participant for a score lower than three times the standard deviation. Two of these participants were removed for problems during both test phases, four for problems during the post-test and finally 16 were removed for problems during the follow-up test. Consequently, we carried out the analysis on a final sample of 142 children, including 32 in the realistic D-VR condition, 44 in the realistic IVR condition, 37 in the anthropomorphic D-VR condition, and 29 in the anthropomorphic IVR condition.

We observed a significant impact of time on the results for the new/old recognition task ($F(1, 138) = 722.556$; $p < .001$, $\eta_p^2 = .840$). As expected, participants performed significantly better during the post-test ($M = .958$, $SD = .175$) compared to the follow-up test ($M = .525$, $SD = .183$). Regarding the impact of immersion on the new/old recognition task results, we found a significant effect ($F(1, 138) = 24.556$; $p < .001$, $\eta_p^2 = .151$). Contrary to our expectations, participants in the D-VR conditions ($M = .798$, $SD = .133$) performed significantly better than participants in the IVR conditions ($M = .688$, $SD = .149$). Additionally, we observed a significant difference in children's performance on the new/old recognition task based on the level of realism in the presentation ($F(1, 138) = 4.544$; $p = .035$, $\eta_p^2 = .032$). Children performed better in the realistic conditions ($M = .757$, $SD = .157$) than in the anthropomorphic conditions ($M = .723$, $SD = .144$). We did not find significant interactions between the levels of immersion and realism of the presentation ($F(1, 138) = .746$; $p = .389$, $\eta_p^2 = .005$) or between time and immersion ($F(1, 138) = .211$; $p = .646$, $\eta_p^2 = .002$). However, we did find a significant interaction between time and realism of the presentation ($F(1, 138) = 4.033$; $p = .047$, $\eta_p^2 = .028$). To explore this interaction further, we computed 2x2 ANOVAs (as reported in the main text, but with a smaller sample size here). During the post-test, realism did not have a significant impact on recognition performance ($F(1, 138) = .432$; $p = .512$, $\eta_p^2 = .003$), but during the follow-up test, participants performed significantly better ($F(1, 138) = 7.926$; $p = .006$, $\eta_p^2 = .054$) in the realistic conditions ($M = .556$, $SD = .196$) than in the anthropomorphic conditions ($M = .488$, $SD = .162$). Finally, we found no significant triple interaction between the time of testing, levels of immersion, and realism of the presentation during the new/old recognition task ($F(1, 138) = 1.491$; $p = .224$, $\eta_p^2 = .011$).

Furthermore, we computed the same model, but with the covariates included. We report this analysis in following tables. Significant results are written in bold.

Repeated measures analysis of covariance for the new/old recognition task, with realism and immersion as between variables, time as a within variable and gender, age, theory of mind (ToM), and verbal abilities as covariates.

Within Effects	$F(1,132)$	P	η_p^2
Day	4.696	0.032	0.034
Day * gender	0.365	0.546	0.003
Day * age	0.474	0.492	0.004
Day * ToM abilities	0.147	0.702	0.001
Day * verbal abilities	2.264	0.135	0.017
Day * immersion	0.037	0.951	<0.001
Day * realism	4.302	0.040	0.032
Day * immersion * realism	0.298	0.586	0.002

Between Effects	$F(1,132)$	P	η_p^2
Gender	0.391	0.533	0.003
Age	3.436	0.066	0.025
ToM abilities	3.540	0.062	0.026
Verbal abilities	4.403	0.038	0.032
Immersion	17.473	<0.001	0.117
Realism	4.724	0.032	0.035
Immersion * realism	0.181	0.671	0.001

Appendix H

We computed a 2 x 2 x 2 ANOVA for mixed designs. The level of immersion and realism of the presentation served as between-subjects variables, while the time of testing (post-test/follow-up test) was considered as within-subjects variable. We had to exclude nine participants because they dropped out or were not able to finish the testing, and two participants for a score lower than three times the standard deviation. Two of these participants were removed for problems during both test phases, one for problems during the post-test and finally eight were removed for problems during the follow-up test. Consequently, we carried out the analysis on a final sample of 153 children, including 32 in the realistic D-VR condition, 44 in the realistic IVR condition, 39 in the anthropomorphic D-VR condition, and 38 in the anthropomorphic IVR condition.

First of all, we did not observe a significant impact of time on the results for the quiz task ($F(1, 149) = 0.102$; $p = .750$, $\eta_p^2 = .001$). Contrary to our expectations, participants during the post-test ($M = 6.05$, $SD = 1.41$) performed equally well as participants during the follow-up test ($M = 6.04$, $SD = 1.40$). We observed a significant impact of immersion on the results for the quiz task ($F(1, 149) = 5.736$; $p = .018$, $\eta_p^2 = .037$). Contrary to our expectations, participants in the D-VR conditions ($M = 6.32$, $SD = 1.07$) performed significantly better than participants in the IVR conditions ($M = 5.81$, $SD = 1.40$). Furthermore, we did not find a significant difference in children's performance on the quiz task depending on the level of realism of the presentation ($F(1, 149) = 0.455$; $p = .501$, $\eta_p^2 = .003$). Children performed equally well in the realistic conditions ($M = 5.97$, $SD = 1.37$) as in the anthropomorphic conditions ($M = 6.12$, $SD = 1.19$). Moreover, we found no significant interaction between the levels of immersion and realism of the presentation ($F(1, 149) = 2.527$; $p = .114$, $\eta_p^2 = .017$), between time and immersion ($F(1, 149) = 2.28$; $p = .138$, $\eta_p^2 = .015$), and between time and realism of the presentation ($F(1, 149) = 0.557$; $p = .457$, $\eta_p^2 = .004$). Finally, we found no significant triple interaction between the time of testing, levels of immersion, and realism of the presentation during the quiz task ($F(1, 149) = 0.626$; $p = .430$, $\eta_p^2 = .004$).

Furthermore, we computed the same model, but with the covariates included. We report this analysis in following tables. Significant results are written in bold.

Repeated measures analysis of covariance for the quiz task, with realism and immersion as between variables, time as a within variable and gender, age, theory of mind (ToM), and verbal abilities as covariates.

Whitin Effects	<i>F</i> (1,145)	<i>P</i>	η_p^2
Day	0.001	0.968	<0.001
Day * gender	0.458	0.500	0.003
Day * age	0.446	0.505	0.003
Day * ToM abilities	0.521	0.471	0.004
Day * verbal abilities	3.691	0.057	0.025
Day * immersion	1.262	0.263	0.009
Day * realism	0.604	0.438	0.004
Day * immersion * realism	1.438	0.232	0.010

Between Effects	<i>F</i> (1,145)	<i>P</i>	η_p^2
Gender	0.979	0.324	0.007
Age	5.124	0.025	0.034
ToM abilities	4.347	0.039	0.029
Verbal abilities	8.561	0.004	0.056
Immersion	2.773	0.098	0.019
Realism	1.025	0.313	0.007
Immersion * realism	2.073	0.152	0.014

Appendix I

Children's responses to the question, "What did the character in the video look like?" according to realism (realistic/anthropomorphic) and immersion (IVR/D-VR).

	<i>Human</i>	<i>Animal</i>
D-VR – Realistic	28	0
IVR – Realistic	42	1
D-VR – Anthropomorphic	0	34
IVR – Anthropomorphic	2	35

Children's responses to the question, "Do you believe what the character told you? Did the character tell the truth?" according to realism (realistic/anthropomorphic) and immersion (IVR/D-VR).

	<i>Yes</i>	<i>No</i>
D-VR – Realistic	28	2
IVR – Realistic	35	3
D-VR – Anthropomorphic	31	2
IVR – Anthropomorphic	36	4

Children's responses to the question, "Can a character like this really exist in our world? Could we meet a character like this in our world?" according to realism (realistic/anthropomorphic) and immersion (IVR/D-VR).

	<i>Yes</i>	<i>No</i>
D-VR – Realistic	10	19
IVR – Realistic	19	22
D-VR – Anthropomorphic	6	30
IVR – Anthropomorphic	9	31

Appendix J

Children’s responses to the question, “Was the presentation just for fun or to learn something?” according to realism (realistic/anthropomorphic) and immersion (IVR/D-VR).

	<i>Fun</i>	<i>Learn</i>
D-VR – Realistic	5	28
IVR – Realistic	5	41
D-VR – Anthropomorphic	4	34
IVR – Anthropomorphic	7	35